

STIC-Biotech/ChemLib

177782  
mey

From: Whiteman, Brian  
Sent: Friday, January 27, 2006 8:28 AM  
To: STIC-Biotech/ChemLib  
Subject: seq search

09/889,874

SEQ ID NO: 22 and 23

1) search us patents and published us patent application databases

Thank you,

Brian Whiteman  
Remsen, 2D14  
mail box 2C18  
Patent Examiner - Art Unit 1635  
United States Patent and Trademark Office  
(571) 272-0764

12/2007  
(SIN)

\*\*\*\*\* Point of Contact:  
\*\*\*\*\* Alexandra Waclawiw  
Searcher: \_\_\_\_\_ Technical Info. Specialist  
Searcher Phone: \_\_\_\_\_ C41 6A02 Tel 303-441  
Date Searcher Picked up: \_\_\_\_\_ 1-30  
Date completed: \_\_\_\_\_ 6 1-30  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_ 8

\*\*\*\*\* Type of Search  
NA# \_\_\_\_\_ AA#: \_\_\_\_\_ 2  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\* Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

**THIS PAGE LEFT BLANK**



CC an engineered bacterium, or a nematocidal protein obtained from such  
 CC bacteria, particularly P13-1f or P14-2f (see AAV95707) can be used to  
 CC control helminths in a human or domesticated animal or for the  
 CC control of plant pathogen nematodes. Also claimed are vectors for  
 CC expressing nematocidal proteins in host cells, and transgenic plants  
 XX SQ Sequence 334 AA;

Query Match 100.0%; Score 1784; DB 3; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-150; Mismatches 0; Indels 0; Gaps 0;  
 Matches 334; Conservative 0;

Db 1 PTLDREDSMDWTGYSTPNTVILETGLDNCNYANGLAMQVINIPTDEGNTUDIDVT 60  
 1 PTLDREDSMDWTGYSTPNTVILETGLDNCNYANGLAMQVINIPTDEGNTUDIDVT 60

Qy 61 LNDNIKIVDQYDSDIDESDGWFTGYNPNBNYNTIPNSQSYLSJKSENQQTQIKRVSCS 120  
 61 LNDNIKIVDQYDSDIDESDGWFTGYNPNBNYNTIPNSQSYLSJKSENQQTQIKRVSCS 120

Db 121 NTSRURTKPSAKVTTSKVVISITONNSNSRVINAINATNFDDERTKETRFNQ 180  
 121 NTSRURTKPSAKVTTSKVVISITONNSNSRVINAINATNFDDERTKETRFNQ 180

Qy 181 SYTSHKSSTNSLYKHTWTPRSLKQLNWRWEDYNGWTWAQSCYKTKGADGGSBTRMIA 240  
 181 SYTSHKSSTNSLYKHTWTPRSLKQLNWRWEDYNGWTWAQSCYKTKGADGGSBTRMIA 240

Db 241 AGS1PPPGNNDGLWDNDTIALSGMAHKSNVDTGINOISFTTRIGKGFSWVNTNISGLRG 300  
 241 AGS1PPPGNNDGLWDNDTIALSGMAHKSNVDTGINOISFTTRIGKGFSWVNTNISGLRG 300

Qy 301 HAVIIIDQYGNKYRLPHAGYENSDPYLSSSIVY 334  
 301 HAVIIIDQYGNKYRLPHAGYENSDPYLSSSIVY 334

Db RESULT 2  
 ABM67282 ID ABM67282 standard; protein; 341 AA.  
 AC ABM67282;  
 XX DT 20-NOV-2003 (first entry)  
 XX DB Photorhabdus luminescens protein sequence #379.  
 XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 KW detection; food; gene expression; plant; animal; microorganism; toxin;  
 KW antibiotic; biopesticide; virulence factor; disease model; plague;  
 KW whooping cough.  
 OS Photorhabdus luminescens.  
 XX PN WO200294867-A2.

XX PD 28-NOV-2002.

XX PR 07-FEB-2002; 2002WO-IB003040.

XX PR 07-FEB-2001; 2001FR-00001659.

XX PR (INSP ) INST PASTEUR.

XX PA (CNRS ) CNRS CENT NAT RECH SCI.

XX PI Duchaude E, Taourit S, Glaser P, Frangoul L, Kunst F, Danchin A;

PI Buchrieser C;

XX DR WPI ; 2003-148459/14.

XX PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides, useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX PR WO200277183-A2.

XX PD 03-OCT-2002.

PS Claim 2; SEQ ID NO 379; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded  
 proteins from Photorhabdus luminescens. The isolated sequences are  
 sources of probes and primers for detecting the genome of P. luminescens  
 and related species; to study polymorphisms; for detection and for  
 detection/amplification of the genes. Antibodies (Ab) raised against the  
 polypeptides encoded by the genes are used for detection/identification  
 of P. luminescens e.g. in foods. The genes, proteins, Ab and cells that  
 carry a gene-containing vector are used to select compounds that  
 modulate, regulate, induce or inhibit expression of the genes in plants,  
 animals or microorganisms other than P. luminescens and are able to alter  
 response or sensitivity to toxins and antibiotics produced by P.  
 luminescens. Cells transformed to express the genes are useful for  
 recombinant production of the proteins, particularly toxins and  
 antibiotics useful as insecticides, bactericides and fungicides. The  
 genes, proteins, vectors containing the genes and Ab are also useful  
 therapeutically to treat microbial infection by bacteria or fungi that  
 are sensitive to P. luminescens-encoded toxins or antibiotics and as  
 biopesticides. Other uses of the genes and the proteins are as virulence  
 factors and for identifying targets of human diseases for which P.  
 luminescens is a model (particularly plague and whooping cough). This  
 sequence represents one of the isolated P. luminescens proteins

XX SQ Sequence 341 AA;

Query Match 9.2%; Score 145.5; DB 6; Length 341;  
 Best Local Similarity 22.9%; Pred. No. 0.00041; Mismatches 137; Indels 53; Gaps 15;  
 Matches 74; Conservative 59; Mismatches 137; Indels 53; Gaps 15;

Db 123 SRURTKPSAKVTTSKVVISITONNSNSRVINAINATNFDDERTKETRFNQ 176  
 123 SRURTKPSAKVTTSKVVISITONNSNSRVINAINATNFDDERTKETRFNQ 176

Db 137 NRTRSTBYLMAKMTPEDTNGKRTTINMSVGDEVFDKSLKLKATAPYANANQHESTNL 196  
 137 NRTRSTBYLMAKMTPEDTNGKRTTINMSVGDEVFDKSLKLKATAPYANANQHESTNL 196

Db 177 FENQSYTSHKSSTNSLYKHTWTPRSLKQLNWRWEDYNGWTWAQSCYKTKGADGGSB 234  
 177 FENQSYTSHKSSTNSLYKHTWTPRSLKQLNWRWEDYNGWTWAQSCYKTKGADGGSB 234

Db 197 FDKEEPT-KSDTHQTINLYRILPYHLIRELEGNDRTNNIYLGKSS 245  
 197 FDKEEPT-KSDTHQTINLYRILPYHLIRELEGNDRTNNIYLGKSS 245

Qy 235 STRMLAASSISPPGNYDQWILDNDIALSGMAHKSYNDTGINOISFTTRIGKGFSWVNTNISGLRG 294  
 235 STRMLAASSISPPGNYDQWILDNDIALSGMAHKSYNDTGINOISFTTRIGKGFSWVNTNISGLRG 294

Db 246 DDRFLTRARVFKRG-T-SYNAARNDMMSGCVMDSYDVTVPDQLAAEVLAHVTCWSWTCY 304  
 246 DDRFLTRARVFKRG-T-SYNAARNDMMSGCVMDSYDVTVPDQLAAEVLAHVTCWSWTCY 304

Qy 295 SGUDRGHAVIIIDQYGNKYRI 315  
 295 SGUDRGHAVIIIDQYGNKYRI 315

Db 305 -VGYHDTIIDLNYGCQHKFRI 325  
 305 -VGYHDTIIDLNYGCQHKFRI 325

RESULT 3  
 ABU42557 ID ABU42557 standard; protein; 892 AA.  
 AC ABU42557;  
 XX DT 19-JUN-2003 (first entry)  
 XX DB Protein encoded by Prokaryotic essential gene #28084.  
 XX KW Antibiotic; prokaryotic essential gene; cell proliferation; drug design.  
 XX OS Staphylococcus epidermidis.  
 XX PN WO200277183-A2.  
 XX PD 03-OCT-2002.

PP 21-MAR-2002; 2002WO-US009107.  
 XX PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-0094893.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 06-MAR-2002; 2002US-0007251.  
 PR (ELIT-) ELITRA PHARM INC.  
 XX PT Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykind JW;  
 XX Wall D, Trwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 DR WPI; 2003-02926/02.  
 DR N-PSDB; AC046427.

XX PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 XX isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 70481; 176pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patient did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 XXX [http://wipo.int/pub/published\\_pct Sequences](http://wipo.int/pub/published_pct Sequences)  
 SQ Sequence 892 AA;

Query Match 7.1%; Score 126; DB 6; Length 892;  
 Basic Local Similarity 21.0%; Pred. No. 0.089; Matches 78; Conservative 51; Mismatches 141; Indels 102; Gaps 16;

Oy 9 SWTGTGTSFNVVPLTGLDNCNYANGLAMIG--VLIINPTPDEGNFDIDVQ---- 60  
 Oy 398 SMTNTDTKNTHTEQ----TIVINPLRYSAKETVNNSINGDDEGSTI-IDBSTITIYKV 451  
 Oy 61 -----LNDNIKIVDVIGSDIDSGDGFWPYTGPNEWT-----IPNS 97  
 Db 452 VEDNONLPDSNRVYDSSYEDVT-NDDYPAQLGNNNDINFRNIDSPVILKISKVDPNK 510  
 Oy 98 QSYSLKSENSTQTOIKRVSCNTSRALKTSKSAV--TTSG-----KVTSI-- 144  
 Db 511 DYTIIQQTWMTTNEY---TGFRTASDNTIAPSTSSGQQGQGDLPBPRTKYKGD 565  
 Oy 145 -----TONSINSSRVVINAIDATNFTDDELRTKETRFENQSYTSHGSNSTNL-Y

RESULT 4  
 ID AD084849  
 ID AD084849 standard; protein; 892 AA.  
 XX AC AD084849;  
 XX DT 29-JUL-2004 (first entry)  
 DB S epidermidis surface anchored LPXTG protein SeqID21.  
 XX LPXTG: cell wall-anchored surface protein; Gram positive bacterium;  
 XX extracellular matrix molecule; sequence database; C-terminal;  
 XX immunoglobulin-like fold region; Ig-like fold region; antibacterial;  
 XX vaccine; gene therapy; infection; medical device; prosthesis;  
 XX premature newborn; AIDS; debilitated cancer; bone marrow transplantation.  
 OS Staphylococcus epidermidis.  
 PN WO2004025416-A2.  
 PD 25-MAR-2004.  
 XX PR 15-SEP-2003; 2003WO-US028789.  
 XX PR 13-SEP-2002; 2002US-0410303P.  
 XX PA (TEXA ) UNIV TEXAS A & M SYSTEM.  
 PA (INHI-) INHIBITEK INC.  
 PA (UAB-) UAB RES FOUND.  
 PI Hutchins JT, Hall A;  
 PI Hutchins JT, Hall A;  
 DR WPI; 2004-315684/29.

PT Identifying LPXTG-containing cell wall-anchored surface proteins from  
 PT Gram positive bacteria, for treating infection caused by the bacteria,  
 PT comprises searching sequence information database for the sequence having  
 PT LPXTG motif.

XX PS Claim 16; SEQ ID NO 21; 96pp; English.

CC This invention relates to a novel method of identifying LPXTG-containing  
 CC cell wall-anchored surface proteins from Gram positive bacteria that bind  
 CC to an extracellular matrix molecule which comprises searching a database  
 CC of sequence information for a putative protein sequence having the LPXTG-  
 CC motif in its C-terminal region and analysing the sequence for the  
 CC presence of one or more Immunoglobulin (Ig)-like fold regions. The  
 CC invention may be useful for the production of compounds with an  
 CC antibacterial activity or for production of a vaccine. In addition the  
 CC disclosed sequences may be useful for gene therapy. The antibody is  
 CC useful for treating or preventing an infection of Gram-positive bacteria  
 CC in a human or animal patient. The method and the proteins are useful in  
 CC generating antibodies for treating and preventing the spread of  
 CC infections of Gram positive bacteria, for interfering with, or inhibiting  
 CC binding interactions by Gram positive bacteria, for monitoring the level

CC antigens in a human or animal patients suspected of containing the  
CC antigens or antibodies, in preventing or reducing infection of medical  
CC devices and prostheses caused by such organisms, and in treating or  
CC preventing infections in highly susceptible groups such as premature  
CC newborns, AIDS and debilitated cancer patients, and bone marrow  
CC transplantation. The present sequence is that of a surface anchored LPXNG  
XX protein identified using the method of the invention.

Sequence 892 AA:

Query Match 7.1%; Score 126; DB 8; Length 892;  
Best Local Similarity 21.0%; Pred. No. 0.089; Mismatches 141; Indels 102; Gaps 16;  
Matches 78; Conservative 51; Mismatches 141; Indels 102; Gaps 16;

Qy 9 SDWGVSTENVLILTGDNCTYANGLNNIG--VIINITPTDDEGNFVDDVT----- 60  
Db 398 SMFVNIDTKNHTVEQ----TIVINPLRYSAKETVNNSISGNGDBGSTI-IDDSIIKVYK 451  
Qy 61 -----INDNIKIVDYGSDSDGSDGFWYTGPBNEMT-----IINS 97  
Db 452 VGDNONLPLPSNRITYDSEYEDVT-NDDYALQGNNDVNINFGNIDSPYIIVKISKYDNK 510  
Qy 98 QSYSLKSENSQIQIKRVUSCSNTSRRTKSFAKV--TTWSG-----KVISI-- 144  
Db 511 DDVYTIQQVTMOTINIEY----TGERRTASYNTIAFSTSQQGQDLPPLPKTYKGD 565  
Qy 145 -----TQNSINSSRVVINAIDATNFTDDELRTKETRPNQSYTSHKSSTNSI-Y 193  
Db 566 YVWEDVDKGIGIONNDNEKEPLSNVLVLTIPDG---TSKSVRTDEBEGKYQFDGLKNGTY 622  
Qy 194 VHTWTIPSLKLQNWWRWY-----NNGWTWAQSCYKTGAGGSSST 236  
Db 623 KITFETP-----EGYPTPLKHSGTNPALDSEGNSWVWTINGQDMTIDSGFYOP 672  
Qy 237 RMLAAGSIFPPGNYDGLWLNDNIALSGMAHKSIKVTDGGINQLSTRIGKGFSPWVNSG 296  
Db 673 KYSIGNYVWYDTNKDGIQGDBEKIGISGV--KVTLKDENGNLISITTTBENG--KYQFDN 727  
Qy 297 LDRAHAAVIIIDQ 308  
Db 728 LNSGNVIVHFDK 739

RESULT 5  
ADS20651 Query Match 7.1%; Score 126; DB 8; Length 892;  
ID ADS20651 Best Local Similarity 21.0%; Pred. No. 0.089; Mismatches 141; Indels 102; Gaps 16;  
XX Matches 78; Conservative 51; Mismatches 141; Indels 102; Gaps 16;  
AC ADS20651; DT 30-DEC-2004 (first entry)  
XX S. epidermidis hyperimmune serum reactive antigen protein - SEQ ID 52.  
XX antigen; antibiotic resistance; antibacterial; vaccine; gene therapy.  
OS Staphylococcus epidermidis.  
XX WO2004087746-A2.  
XX PD 14-OCT-2004.  
XX PP 31-MAR-2004; 2004WO-EP003398.  
XX PR 31-MAR-2003; 2003EP-00450078.  
XX PA (INTE-) INTERCELL AG.

RESULT 6  
ABP40469 Query Match 7.1%; Score 126; DB 8; Length 930;  
ID ABP40469 Best Local Similarity 21.0%; Pred. No. 0.089; Mismatches 141; Indels 102; Gaps 16;  
AC ABP40469; DT 24-JUL-2002 (first entry)  
XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5314.  
XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
XX N-PSDB; ADS20620.  
OS New nucleic acid molecules encoding hyperimmune serum reactive antigens

from *Staphylococcus epidermidis*, useful for diagnosing, preventing or  
treating *S. epidermidis* infections.

XX Claim 11; SEQ ID NO 52; 196pp; English.

The invention relates to a novel isolated nucleic acid molecule encoding  
a hyperimmune serum reactive antigen or its fragment. *Staphylococci* are  
commonly associated with human disease. Both *Staphylococcus epidermidis*  
and *Staphylococcus aureus* have become resistant to many commonly used  
antibiotics, most importantly methicillin (MRSA) and vancomycin (VISA).  
Drug resistance is an increasingly important public health concern and  
novel therapies to combat *staphylococci* infection must be developed in  
preparation for a time when such infections may be untreatable by  
antibiotics. The molecules of the invention demonstrate antibacterial  
activity and may be useful for manufacturing a medicament, such as a  
vaccine, for treating or preventing *S. epidermidis* infections possibly  
via gene therapy. The antigen or its fragment may also be used for  
generating an anticancer peptide binding to the hyperimmune serum  
reactive antigen or fragment, for manufacturing a functional nucleic  
acid, such as an aptamer or spiegelmer and for manufacturing a functional  
ribonucleic acid, such as a ribozyme, antisense nucleic acid or siRNA  
(short interfering RNA). The current sequence is that of a *Staphylococcus*  
*epidermidis* hyperimmune serum reactive antigen protein of the invention.

Sequence 892 AA:

Query Match 7.1%; Score 126; DB 8; Length 892;  
Best Local Similarity 21.0%; Pred. No. 0.089; Mismatches 141; Indels 102; Gaps 16;  
Matches 78; Conservative 51; Mismatches 141; Indels 102; Gaps 16;

Qy 9 SDWGVSTENVLILTGDNCTYANGLNNIG--VIINITPTDDEGNFVDDVT----- 60  
Db 398 SMFVNIDTKNHTVEQ----TIVINPLRYSAKETVNNSISGNGDBGSTI-IDDSIIKVYK 451  
Qy 61 -----INDNIKIVDYGSDSDGSDGFWYTGPBNEMT-----IINS 97  
Db 452 VGDNONLPLPSNRITYDSEYEDVT-NDDYALQGNNDVNINFGNIDSPYIIVKISKYDNK 510  
Qy 98 QSYSLKSENSQIQIKRVUSCSNTSRRTKSFAKV--TTWSG-----KVISI-- 144  
Db 511 DDVYTIQQVTMOTINIEY----TGERRTASYNTIAFSTSQQGQDLPPLPKTYKGD 565  
Qy 145 -----TQNSINSSRVVINAIDATNFTDDELRTKETRPNQSYTSHKSSTNSI-Y 193  
Db 566 YVWEDVDKGIGIONNDNEKEPLSNVLVLTIPDG---TSKSVRTDEBEGKYQFDGLKNGTY 622  
Qy 194 VHTWTIPSLKLQNWWRWY-----NNGWTWAQSCYKTGAGGSSST 236  
Db 623 KITFETP-----EGYPTPLKHSGTNPALDSEGNSWVWTINGQDMTIDSGFYOP 672  
Qy 237 RMLAAGSIFPPGNYDGLWLNDNIALSGMAHKSIKVTDGGINQLSTRIGKGFSPWVNSG 296  
Db 673 KYSIGNYVWYDTNKDGIQGDBEKIGISGV--KVTLKDENGNLISITTTBENG--KYQFDN 727  
Qy 297 LDRAHAAVIIIDQ 308  
Db 728 LNSGNVIVHFDK 739

US6380370-B1.  
PN XX  
PD XX  
XX 30-APR-2002.  
XX PF 13-AUG-1998; 98US-00134001.  
PR 14-AUG-1997; 97US-0055779P.  
PR 08-NOV-1997; 97US-0064964P.  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX PT Doutette-Stamm LA, Bush D;  
XX DR WPI; 2002-381255/41.  
DR N/PDB; ABN93014.  
XX PT Novel isolated nucleic acid encoding a *Staphylococcus epidermidis*  
PT polypeptide, useful for diagnosing and treating bacterial infections.  
XX PS Disclosure; SEQ ID NO 5314; 267pp; English.  
XX ABN930538 to ABN93374 represent *Staphylococcus epidermidis* open reading  
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
CC given in ABP512 to ABP37960. The *S. epidermidis* sequences have  
CC antibacterial activity and can be used in gene therapy. The sequences can  
CC also be used in the diagnosis and treatment of bacterial infections,  
particularly *S. epidermidis* infections. The sequences can be used to  
CC screen for compounds able to interfere with the *S. epidermidis* life cycle  
CC or inhibit *S. epidermidis* infection. N.B. The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from the USPTO web site.  
Sequence 930 AA;

04-NOV-2004 (first entry)  
**Staphylococcus epidermidis polypeptide seqid 5309.**  
antibacterial; vaccine; antisense therapy; **Staphylococcus epidermidis**; recombinant expression vector; infection; computer readable medium; computer based system.  
**Staphylococcus epidermidis:**  
US2004147734-A1.  
29-JUL-2004.  
01-DBC-2003; 2003US-00724972.  
08-NOV-1997; 97US-0064964P.  
13-AUG-1998; 98US-00134001.  
29-NOV-1999; 99US-00450969.  
(DOUC/) DOUCETTE-STAMM L.  
(BUSH/) BUSH D.  
Doucette-Stamm L, Bush D;  
WPI; 2004-580138/55.  
N-PSDB; ADS02242.  
New isolated polypeptide and encoding nucleic acid derived from **Staphylococcus epidermidis**, useful for diagnosing, preventing and/or treating an **S. epidermidis** bacterial infection.  
Claim 17; SEQ ID NO 5309; 741PP; English.

The invention describes an isolated nucleic acid comprising a nucleotide sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO: 1-3772) and encoding an *Staphylococcus epidermidis* polypeptide with any of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as given in the specification. Also described are: a recombinant expression vector; a cell comprising a recombinant expression vector of (1); producing an *S. epidermidis* polypeptide; an isolated nucleic acid comprising a nucleotide sequence of at least 8 nucleotides in length; a vaccine composition for prevention or treatment of an *S. epidermidis* infection, comprising a nucleic acid cited above and a carrier; treating a subject for *S. epidermidis* infection, a recombinant or substantially pure preparation of an *S. epidermidis* polypeptide or its fragment; a vaccine composition for prevention or treatment of an *S. epidermidis* infection; detecting the presence of a *Staphylococcus* nucleic acid in a sample; a computer readable medium having recorded in it the nucleotide sequences with SEQ ID NO: 1-3772 or its fragments; a computer based system for identifying fragments of the *Staphylococcus* genome of commercial importance; a computer based system for identifying fragments of the *Staphylococcus* plasmids of commercial importance; identifying commercially important nucleic acid fragments of the *Staphylococcus* genome and/or plasmids; and identifying an expression modulating fragment of the *Staphylococcus* genome and/or plasmids. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of an *Staphylococcal* epidermidis bacterial infection. This is the amino acid sequence of a *S. epidermidis* protein of the invention.

PS Claim 8; Fig 3; 104pp; English.

XX Isolated *Staphylococcus* Sdr cell wall proteins which bind both soluble

CC and immobilized fibrinogen are useful for treating or preventing

CC coagulase-negative *staphylococcal* infection such as *Scepticemia*, in

CC osteomylelitis or endocarditis, and for inducing immune responses in

patients. The cell wall proteins are also useful for reducing coagulase-

negative *staphylococci* infection of indwelling medical devices such as

vascular grafts, vascular stents, intravenous catheters, artificial heart

valves and cardiac assist devices. The cell wall associated proteins are

able to inhibit *Staphylococcal* adhesion to immobilized extracellular

CC matrix or host cells present on the surface of implanted biomaterials

XX

Sequence 991 AA;

Query Match 7.1%; Score 126; DB 3; Length 991;

Best Local Similarity 21.0%; Pred. No. 0..1.; Gaps 16;

Matches 78; Conservative 51; Mismatches 141; Indels 102;

XX

AC AAY83171;

XX

DT 24-JUL-2000 (first entry)

XX

DE Cell wall protein SdrG.

XX

Key Location/Qualifiers

KW SdrP; SdrG; SdrH; coagulase negative; *staphylococcus*; septicemia;

KW intravenous catheter; heart valve; cardiac.

KW Osteomyelitis; endocarditis; immune response; vaccine; graft; stent;

KW *Staphylococcus* sp.

XX

FT Misc-difference 14 /note= "Position encoded by TAG stop codon"

FT Misc-difference 33 /note= "Position encoded by TGA stop codon"

FT Misc-difference 964 /note= "Position encoded by TAA stop codon"

FT Misc-difference 980 /note= "Position encoded by TAG stop codon"

FT Misc-difference 989 /note= "Position encoded by TAA stop codon"

FT Misc-difference 989 /note= "Position encoded by TAA stop codon"

XX

PN WO20012699-A1.

XX

PD 09-MAR-2000.

XX

PP 31-AUG-1999; 99NO-US01972B.

XX

PR 31-AUG-1998; 98US-0098443P.

PR 25-JAN-1999; 99US-017719P.

XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

PA (TEKA) UNIV TEXAS A & M SYSTEM.

XX

PT Foster TJ, Hook M, Davis S, Hartford O, McCrea K, Ni Eilhin D;

XX

DR WPI, 2000-25637/22.

DR N-PADB; AAY93534.

XX

PT Recombinant or synthetic proteins from coagulase-negative *staphylococci*

PT useful for prevention, treatment and diagnosis of *staphylococcal*

PT infections bind soluble and immobilized fibrinogen.

XX

PS Claim 8; Fig 3; 104pp; English.

XX Isolated *Staphylococcus* Sdr cell wall proteins which bind both soluble

CC and immobilized fibrinogen are useful for treating or preventing

CC coagulase-negative *staphylococcal* infection such as *Scepticemia*, in

CC osteomylelitis or endocarditis, and for inducing immune responses in

patients. The cell wall proteins are also useful for reducing coagulase-

negative *staphylococci* infection of indwelling medical devices such as

vascular grafts, vascular stents, intravenous catheters, artificial heart

valves and cardiac assist devices. The cell wall associated proteins are

able to inhibit *Staphylococcal* adhesion to immobilized extracellular

CC matrix or host cells present on the surface of implanted biomaterials

XX

Sequence 991 AA;

Query Match 7.1%; Score 126; DB 3; Length 991;

Best Local Similarity 21.0%; Pred. No. 0..1.; Gaps 16;

Matches 78; Conservative 51; Mismatches 141; Indels 102;

XX

AC AAY83171;

XX

DT 24-JUL-2000 (first entry)

XX

DE Cell wall protein SdrG.

XX

Key Location/Qualifiers

KW SdrP; SdrG; SdrH; coagulase negative; *staphylococcus*; septicemia;

KW intravenous catheter; heart valve; cardiac.

KW Osteomyelitis; endocarditis; immune response; vaccine; graft; stent;

KW *Staphylococcus* sp.

XX

FT Misc-difference 14 /note= "Position encoded by TAG stop codon"

FT Misc-difference 33 /note= "Position encoded by TGA stop codon"

FT Misc-difference 964 /note= "Position encoded by TAA stop codon"

FT Misc-difference 980 /note= "Position encoded by TAG stop codon"

FT Misc-difference 989 /note= "Position encoded by TAA stop codon"

FT Misc-difference 989 /note= "Position encoded by TAA stop codon"

XX

PN WO20012699-A1.

XX

PD 09-MAR-2000.

XX

PP 31-AUG-1999; 99NO-US01972B.

XX

PR 31-AUG-1998; 98US-0098443P.

PR 25-JAN-1999; 99US-017719P.

XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

PA (TEKA) UNIV TEXAS A & M SYSTEM.

XX

PT Foster TJ, Hook M, Davis S, Hartford O, McCrea K, Ni Eilhin D;

XX

DR WPI, 2000-25637/22.

DR N-PADB; AAY93534.

XX

PT Recombinant or synthetic proteins from coagulase-negative *staphylococci*

PT useful for prevention, treatment and diagnosis of *staphylococcal*

PT infections bind soluble and immobilized fibrinogen.

XX

PS Claim 8; Fig 3; 104pp; English.

XX Isolated *Staphylococcus* Sdr cell wall proteins which bind both soluble

CC and immobilized fibrinogen are useful for treating or preventing

CC coagulase-negative *staphylococcal* infection such as *Scepticemia*, in

CC osteomylelitis or endocarditis, and for inducing immune responses in

patients. The cell wall proteins are also useful for reducing coagulase-

negative *staphylococci* infection of indwelling medical devices such as

vascular grafts, vascular stents, intravenous catheters, artificial heart

valves and cardiac assist devices. The cell wall associated proteins are

able to inhibit *Staphylococcal* adhesion to immobilized extracellular

CC matrix or host cells present on the surface of implanted biomaterials

XX

Sequence 991 AA;

Query Match 7.1%; Score 126; DB 3; Length 991;

Best Local Similarity 21.0%; Pred. No. 0..1.; Gaps 16;

Matches 78; Conservative 51; Mismatches 141; Indels 102;

XX

AC AAY83171;

XX

DT 24-JUL-2000 (first entry)

XX

DE Cell wall protein SdrG.

XX

Key Location/Qualifiers

KW SdrP; SdrG; SdrH; coagulase negative; *staphylococcus*; septicemia;

KW intravenous catheter; heart valve; cardiac.

KW Osteomyelitis; endocarditis; immune response; vaccine; graft; stent;

KW *Staphylococcus* sp.

XX

FT Misc-difference 14 /note= "Encoded by in-frame stop codon TAG"

FT Misc-difference 33 /note= "Encoded by in-frame stop codon TGA"

FT Misc-difference 964 /note= "Encoded by in-frame stop codon TAA"



CC block the adherence of bacteria) and for diagnosis. (Updated on 17-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 1092 AA;

Query Match 7.0%; Score 124; DB 2; Length 1092;  
 Best Local Similarity 21.0%; Pred. No. 0.18; Mismatches 142; Indels 102; Gaps 16;  
 Matches 78; Conservative 50; Mismatches 142; Indels 102; Gaps 16;

QY 9 SDWTGVSTPNVILETGGLDNCNTYANGLMMIG--VIINITPTDDDEGNFVDDVT---- 60  
 Db 439 SMFTNIDTKHTEQ----TIVINPLRYSAKETVNWNISGNGSBERGSTI-IDDSTIKWYK 492

QY 61 -----LNNDNIKLKVUDYIYGSDIDGSDGWFYTGPNEYNT----IPNS 97  
 Db 493 VGDQNQLPDSNRIDYSEVEDVT-NDDYAQLGNNNDVINGFGNIDSPYIIVKISKYDPK 551

QY 98 QSYSLIKSENSQTOIQKRVCSNTSRKTSKSAV--TTSG----KVSI-- 144  
 Db 607 YVWEDVDKDGITNDNEKPLSNLVLTVPDG--TSKSVRTEDGKXQFDLKGITY 663

QY 194 VHTWTPRSILKQLQWRWEDY-----NNGWTWAQSCYKTKGADGGSST 236  
 Db 664 KITFETP-----EGYPTPLKHSCTNPALDSEGNSVWVTINGQDMTLDGFQOTP 713

QY 237 RMLAAGSTFPPGNYDGLWDNDIALSGMAHKSYNDTGGINQSLSPTRIGKGFSWVNMSG 296  
 Db 714 KYSLGNVWYDNTNKDGIGQDDEKGISGV--KVTLKDENGNISTTTDENG--KYQFDN 768

QY 297 LDRGHAVITIDQ 308  
 Db 769 LNSGNVYVHDK 780

RESULT 11  
 ABM79019  
 ID ABM79019 standard; protein; 1092 AA.

AC ABM79019;  
 XX  
 DT 15-JAN-2004 (first entry)

XX DE Staphylococcus epidermidis polypeptide.  
 XX  
 KW Infection; antibacterial; vaccine.

XX OS Staphylococcus epidermidis.  
 XX PN WO2003076470-A1.

XX PD 18-SEP-2003.  
 XX PP 05-MAR-2003; 2003WO-US006415.

PR 05-MAR-2002; 2002US-0361324P.

XX PA (INH-) INHIBITEK INC.  
 PA (TEXA ) UNIV TEXAS A & M SYSTEM.

XX PI Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;  
 PI Robbins J, Vernachio J, Bowden MG;  
 XX WPI; 2003-722324/68.

XX PT New antibody recognizing a Staphylococcus epidermidis protein comprising  
 PT SdrG N1N3, SdrG N2N3 or SdrGTR2 useful for preparing a composition for  
 PT treating or preventing a coagulase-negative Staphylococcal infection. for  
 XX Claim 16; Page 36-37; 78pp; English.

XK The present sequence comprises the protein sequence of a polypeptide of a  
 CC coagulase-negative *Staphylococcus epidermidis*. A claimed monoclonal  
 CC antibody recognises this protein and is used in a claimed method of  
 CC treating or preventing a coagulase-negative *Staphylococcal* infection in a  
 CC human or animal, e.g. a nosocomial coagulase-negative *staphylococcal*  
 CC infection in low birth weight infants

XK Sequence 1092 AA;

Query Match 7.0%; Score 124; DB 7; Length 1092;  
 Best Local Similarity 21.0%; Pred. No. 0.18; Mismatches 142; Indels 102; Gaps 16;  
 Matches 78; Conservative 50; Mismatches 142; Indels 102; Gaps 16;

QY 9 SDWTGVSTPNVILETGGLDNCNTYANGLMMIG--VIINITPTDDDEGNFVDDVT---- 60  
 Db 439 SMFTNIDTKHTEQ----TIVINPLRYSAKETVNWNISGNGSBERGSTI-IDDSTIKWYK 492

QY 61 -----LNNDNIKLKVUDYIYGSDIDGSDGWFYTGPNEYNT----IPNS 97  
 Db 493 VGDQNQLPDSNRIDYSEVEDVT-NDDYAQLGNNNDVINGFGNIDSPYIIVKISKYDPK 551

QY 98 QSYSLIKSENSQTOIQKRVCSNTSRKTSKSAV--TTSG----KVSI-- 144  
 Db 607 YVWEDVDKDGITNDNEKPLSNLVLTVPDG--TSKSVRTEDGKXQFDLKGITY 663

QY 194 VHTWTPRSILKQLQWRWEDY-----NNGWTWAQSCYKTKGADGGSST 236  
 Db 664 KITFETP-----EGYPTPLKHSCTNPALDSEGNSVWVTINGQDMTLDGFQOTP 713

QY 237 RMLAAGSTFPPGNYDGLWDNDIALSGMAHKSYNDTGGINQSLSPTRIGKGFSWVNMSG 296  
 Db 714 KYSLGNVWYDNTNKDGIGQDDEKGISGV--KVTLKDENGNISTTTDENG--KYQFDN 768

QY 297 LDRGHAVITIDQ 308  
 Db 769 LNSGNVYVHDK 780

RESULT 12  
 ADV16755  
 ID ADV16755 standard; protein; 893 AA.

AC ADV16755;  
 XX  
 DT 24-FEB-2005 (first entry)

XX DE E. faecalis V583 hyperimmune serum reactive antigen protein - SEQ ID 450.

XX KW antigen; antibacterial; vaccine; enterococcus infection; infection;  
 KW pharyngitis; impetigo; rheumatic fever; antipyretic; antirheumatic;  
 KW immunosuppressive; sepsis.

XX OS Enterococcus faecalis V583.

XX PR Key Location/Qualifiers

XX PT Region 6..30  
 XX PT /note= "Serum reactive epitope, claimed"  
 XX Region 36..42 "Serum reactive epitope, claimed"  
 XX PT /note= "Serum reactive epitope, claimed"  
 XX Region 143..157 /note= "Serum reactive epitope, claimed"  
 XX PT Region 176..197 /note= "Serum reactive epitope, claimed"  
 XX Region 205..209 /note= "Serum reactive epitope, claimed"  
 XX Region 216..233 /note= "Serum reactive epitope, claimed"  
 XX PT Region 226..269 /note= "Serum reactive epitope, claimed"



XX  
PD 03-OCT-2002.  
XX  
PP 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00915242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923D.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PT Wang L, Zamudio C, Malone C, Haselbeck R, Olsen KJ, Zytkind JW;  
PT Wall D, Trawick JD, Carr GD, Yamamoto R, Forsyth RA, Xu HH;  
XX DR WPI; 2003-029926/02.  
DR N-PSDB; ACA46390.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 7044; 1766pp; English.  
XX



**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 30, 2006, 09:42:59 ; Search time 7.48879 seconds  
3687.335 Million cell updates/sec

Title: US-09-889-874a-22

perfect score: 1784

Sequence: FTLRDMSDWGVTSPFNVI.....ILFHAGYENSDPYLSSSTVY 334

Scoring table: BLOSUM62

Gapop 10.0 , Gapext: 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5 COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/6 COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/H COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/PCTUS COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/RE COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/backfleis1.pep:\*

Pred. No. 19 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and 18 is derived by analysis of the total score distribution.

## SUMMARIES

\*  
#

Result No.	Score	Query Match	Length	DB ID	Description
1	126	7.1	930	2	US-09-134-001C-5314
2	126	7.1	930	2	US-09-134-001C-5314
3	124	7.0	1092	2	US-09-147-405B-15
4	119	6.7	899	2	US-09-134-000-4697
5	115.5	6.5	1411	2	US-10-080-505-17
6	114	6.4	360	2	US-09-107-433-5178
7	114	6.4	376	2	US-09-109-582-110-5138
8	112.5	6.3	1436	2	US-10-080-505-13
9	107.5	6.0	324	2	US-09-830-230A-418
10	107.5	6.0	343	2	US-09-830-230A-417
11	106	5.9	259	2	US-09-253-645-5
12	105	5.9	1430	2	US-09-008-172-2
13	105	5.9	1430	2	US-09-210-361-6
14	105	5.9	1430	2	US-09-740-274-6
15	104	5.8	1306	2	US-09-538-092-330
16	104	5.8	1589	2	US-09-543-681A-4998
17	103.5	5.8	582	2	US-09-147-405B-13
18	103.5	5.8	593	2	US-09-147-405B-11
19	103	5.8	494	2	US-09-543-681A-4773
20	103	5.8	607	1	US-08-403-995-6
21	103	5.8	607	2	US-08-685-467-6
22	103	5.8	607	2	US-08-913-942-6
23	103	5.8	607	2	US-09-684-707-6
24	103	5.8	1335	2	US-10-080-505-7
25	103	5.8	1912	1	US-08-409-954-4
26	103	5.8	1912	1	US-08-683-467-4
27	103	5.8	2353	2	US-09-377-155-33

## ALIGNMENTS

RESULT 1  
US-09-134-001C-5314

; Sequence 5314, Application US/09134001C  
; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTO-007

; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR PUBLISHING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5314

; LENGTH: 930

; TYPER: PTM  
; ORGANISM: Staphylococcus epidermidis

; US-09-134-001C-5314

Query Match 7.1%; Score 126; DB 2; Length 930;

Best Local Similarity 21.0%; Pred. No. 0.0057; Mismatches 51; Indels 102; Gaps 16;

Matches 78; Conservative 51; Sequence 10, Appl

Sequence 15, Appl

Sequence 497, Appl

Sequence 17, Appl

Sequence 5178, Appl

Sequence 5138, Appl

Sequence 13, Appl

Sequence 418, Appl

Sequence 417, Appl

Sequence 5, Appl

Sequence 2, Appl

Sequence 6, Appl

Sequence 330, Appl

Sequence 4998, Appl

Sequence 13, Appl

Sequence 11, Appl

Sequence 4773, Appl

Sequence 6, Appl

Sequence 6,

RESULT 2  
US-09-386-962C-10  
Sequence 10, Application US/09386962C  
; Patient No. 663473  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy  
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCUS  
; FILE REFERENCE: P06335US2/BAS  
; CURRENT APPLICATION NUMBER: US/09/386, 962C  
; CURRENT FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/098, 443  
; PRIOR FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: SE 9602496-3  
; PRIOR FILING DATE: 1996-06-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 1092  
; TYPE: PRT  
; ORGANISM: *Staphylococcus epidermidis*  
US-09-386-962C-10

Query Match 7.1%; Score 126; DB 2; Length 930;  
Best Local Similarity 21.0%; Pred. No. 0.0057; Mismatches 141; Indels 102; Gaps 16;  
Matches 78; Conservative 51; Mismatches 141; Indels 102; Gaps 16;

QY 9 SDWTGCVSTENVILEGLDNCONIYANGLNMIG--VIINIPPTDDEGFNFVFDIDVT---- 60  
Db 436 SMFNPNTDPTKNTHTVEQ----TIVINPLRVSAKEETVNUNSGNGBGSTI-IDDSTIKYK 489

QY 61 -----LNDNIKIVDYGSDGSDGFWYTGNINEYNT-----IENS 97  
490 VGDNONLPPSNRKYDSEVEDVT-NDDYAQLGNNDVNINFQHIDSPYIIVKISKYDNK 548

QY 98 QSYSLIKSENQTOIKRVSCTSNTSLRKTFSAKV-TTSC-----KVIST-- 144  
Db 549 DYYTTIQQTWMTQTTNEY---TGEFRATASYNTIAESTSSSQGQGDLPBPKTYKGD 603

QY 145 -----TONSINSRRVINAIDATNFTDLRTRKETREFENQSYTSHKESTNSL-Y 193  
Db 604 YVWEDVDVKODGINQNTDNEKPLSNLVLTYPDG---TSKSVRIDEKGKQFDLGKNGLY 660

QY 194 VHTWTIPLSILQWWRWEDY-----NNGWTAQSCYKTAGDGGSIST 236  
Db 661 KITFETP-----EGYPTPLKHSTCNPALDSEGNSVWVTINGQDDMTIDSGFYKTP 710

QY 237 RWLALGSTFPFGNYDGLWLDNDTALSQMAHKSNVDTQINQSFTRIGKGFSWVNISG 296

Db 711 KYSIGNYVWDTNKDGQIQQGDDBKGISGV--KVTLKDENGNIISTTIDENG--KIQDFN 765

QY 297 LDRGHAVITIDQ 308  
Db 766 LNSGNYIVHFDK 777

RESULT 3  
US-09-147-405B-15  
Sequence 15, Application US/09147405B  
; Patient No. 6733758  
; GENERAL INFORMATION:  
; APPLICANT: Guss, Bengt  
; APPLICANT: Nilsson, Martin  
; APPLICANT: Frykberg, Lars  
; APPLICANT: Flock, Jan-Ingar  
; TITLE OF INVENTION: Fibronogen Binding Protein Originating from  
; FILE REFERENCE: Coagulase-Negative *Staphylococcus*

FILE REFERENCE: guss 09/147405  
TITLE OF INVENTION: Coagulase-Negative *Staphylococcus*

RESULT 4  
US-09-134-000C-4697  
Sequence 4697, Application US/09134000C  
; Patient No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Dorette-Stamm et al  
; TITLE OF INVENTION: NUCIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134, 000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055, 778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4697  
; LENGTH: 899  
; TYPE: PRT  
; ORGANISM: *Enterococcus faecalis*  
US-09-134-000C-4697

Query Match 6.7%; Score 119; DB 2; Length 899;  
Best Local Similarity 19.4%; Pred. No. 0.025; Mismatches 119; Indels 130; Gaps 20;  
Matches 74; Conservative 58; Mismatches 119; Indels 130; Gaps 20;

QY 21 LFTGLDN-----CNIYANGLNMIGVINIPTDDEGFNFVFDID---DVLNDNIKIVD 69

RESULT 5  
US-10-080-505-17  
Sequence 17, Application US/10080505  
; Patent No. 6676948  
; GENERAL INFORMATION:  
; APPLICANT: St. Geme, Joseph W.  
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS  
; FILE REFERENCE: A-59941-/RFT/DC/DHR  
; CURRENT APPLICATION NUMBER: US/10/080,505  
; CURRENT FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: US 08/236,791  
; PRIOR FILING DATE: 1994-10-25  
; PRIOR APPLICATION NUMBER: US 09/839,996  
; PRIOR FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 1411  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
; US-10-080-505-17

Query Match 6.5%; Score 115.5; DB 2; Length 1411;  
Best Local Similarity 20.8%; Pred. No. 0.11; Mismatches 131; Indels 145; Gaps 21; Matches 85; Conservative 48;

Qy 4 REDMSMDWTGVS-----TENVILIGLNDNCNTIYANGLAMIGVINIT----- 45  
Db 554 RDAFNGWEGDKDDTKNTERGRNVNPL---NKDNFHFLSGTNLKG---NITDGGSTL 607  
Qy 46 ---PTDDEGNFVD-----IDDVTLNDNIKIVDY-IDG-----SDIG 78  
Db 608 PSGRPTPHAYNHLRNLEGRPKKEVWIDDWMNRTPKENFQIKGGTVVSNVSSIEG 667  
Qy 79 SDGMFYTGPNP--EYNTIPNSQSYSLKSENQTOIKRVCSNTSRKTSFSAKVTT 137  
Db 668 N-WTISNNANATFGVVPENOQNTCTRSIDWTGLT-----TCKIVNLTDKVKIDSIPTQ 719  
Qy 138 SGKUITSITONS-----INSSRVWIN-----AIDATN- 163  
Db 720 INGSINLTNNATVNHLAKLGNVTLHMSOFTLSNNATQGTNQLSNHNATDNTN 779  
Qy 164 -----FTDDEBLRTKETREPNQSYTHSKESTNSLYVHTWTPISLKLQWRWEDYNWT 218  
Db 780 NGNVHLTDSAQFSUKSHFSHQ-IQGDKOTTVLLENATWTMPSDTQNL---TIANSTV 835

RESULT 5  
US-09-107-433-5178  
Sequence 5178, Application US/09107433  
; Patent No. 6800744  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS  
; NUMBER OF SEQUENCES: 5206  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02454  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: <Unknown>  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: <Unknown>  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,433  
; FILING DATE: 30-Jun-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 60/ 085131  
; FILING DATE: May 12, 1998  
; APPLICATION NUMBER: 60/051553  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arinello, Pamela Deneka  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTCC-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5178:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 360 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEtical: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1.. 360  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5178:  
; US-09-107-433-5178

Query Match 6.4%; Score 114; DB 2; Length 360;  
Best Local Similarity 25.6%; Pred. No. 0.019; Mismatches 73; Indels 28; Gaps 8; Matches 46; Conservative 33;

Qy 31 YANGLNMGIVIN-ITPDDEGNFVDIDVTLNDNIKIVDY-IDG-----SDIG 78  
Db 104 YKNAHOALSVLNRGAATPITENDSUVDEVKGANDTLSAQVAMQADLVLVLTVDG 163  
Qy 79 SDGMFYTGPNP--EYNTIPNSQSYSLKSENQTOIKRVCSNTSRKTSFSAKVTT 137  
Db 164 ---LYTCNP--NSDPRAKRRIETINREIDMAGGAGSSNGTSGMLTKAAATATE 216



RESULT 10

US-09-830-230A-417

; sequence 417, Application US/09830230A

; Patent No. 690293

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.

; TITLE OF INVENTION: Lyme Disease Vaccines

; FILE REFERENCE: PBA41IUS

; CURRENT APPLICATION NUMBER: US/09/0930,230A

; PRIOR FILING DATE: 2001-09-27

; PRIOR APPLICATION NUMBER: PCT/US98/12718

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/057,483

; PRIOR FILING DATE: 1997-09-03

; PRIOR APPLICATION NUMBER: 60/053,344

; PRIOR FILING DATE: 1997-07-22

; PRIOR APPLICATION NUMBER: 60/053,377

; PRIOR FILING DATE: 1997-07-22

; PRIOR APPLICATION NUMBER: 60/050,359

; PRIOR FILING DATE: 1997-06-20

; NUMBER OF SEQ ID NOS: 755

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 417

; LENGTH: 343

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-830-230A-417

Query Match 6.0%; Score 107.5; DB 2; Length 343;

Best Local Similarity 18.4%; Pred. No. 0.073; Mismatches 80; Indels 77; Gaps 9; Matches 44; Conservative 38; Mismatches 80; Indels 77; Gaps 9;

Qy 14 VSTENVILLETGLDNCNIVANGLNMGIVNINITPDPDEGNFVDDVLTNDNIKIVDYG 73

Db 141 INYKTFILNGIDM-----NEEKSYFFNLNSD--VQDY--- 177

Qy 74 SDIDGSDGMPTGNPNEYNTIP-----NSQSYLLKSENQTO 112

Db 178 ----DESFYY--KKFLISIPRAHLKIDSRSRDFNVTKINYTFNPENPVYRVLGDIQ 229

Qy 113 KRTVSCSNTSRL----RTKSFSAKVTTGKVVISITONINSNSRVVINAIDATNFDD 167

Db 230 VKNFVLSGNTSKLNRDKQFFIQSDWOKGK----SNSINNSFLTMIRLGGRRN 284

Qy 168 ELRTKETRENGQSYTHGSSTSNSLYVHW----TIPRSIKLQNRWEDTNGTW 219

Db 285 GIOPAKHLEADDSDISDLYESRGDHIEHWYFYKRIVPKD----PEINGNW 315

RESULT 11

US-09-251-645-5

; Sequence 5, Application US/09251645

; Patent No. 612602

; GENERAL INFORMATION:

; APPLICANT: Kramer, Scott E.

; TITLE OF INVENTION: Substitutes for Modified Starch and

; FILE REFERENCE: 058D

; CURRENT APPLICATION NUMBER: US/09/008,172

; EARLIER APPLICATION NUMBER: 08/482,711

; EARLIER FILING DATE: 1995-06-07

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO: 2

; LENGTH: 1430

; TYPE: PRT

; ORGANISM: Streptococcus mutans

US-09-251-645-5

Query Match 5.9%; Score 105; DB 2; Length 1430;

Best Local Similarity 22.5%; Pred. No. 1.1; Mismatches 40; Indels 110; Gaps 19; Matches 82; Conservative 40; Mismatches 133; Indels 110; Gaps 19;

Qy 5 EDMSMDWTGVSTENVILLETGLDNCN--TYANGLNMGIVN--ITPDDEGNF---VD 55

Db 411 EDNSS----GGYDFLLANDISDNPVQAOLNWLHMLNYGSIVANDPEANFDGVFD 465

Qy 56 IDDVTLNDNIKIV-DYDGS--DIDGS----DGWFTGPNPHEYN-----TI 94

Db 466 AVDNVADLQLQASDYLKAHYGVDKSEKNAHNHSILEAW--SDNDPQYNKOTKGQPLI 523

Qy 95 PNSQSYSL-----LXSENSQITQIKGVSCNTSKRTFSAKVTTGKV 147

Db 524 DNTRLSSLYALTRPLEKDASNNEITRGLPEVITNSLNNRSEAEGKNSERMANYIFRAH 583

Qy 148 SINSSRVINAL-----DATNFTDDELRT----TKEFENSYT----- 183

Db 504 DSEVQTVIAKLIKAQINPKTDSLTFDLEKAFKLYMEDHQAKKKTQSNIPTAYLM 643

Qy 184 -SHKSSTSNSLY-----VHTWTPR-----SLKLQNWRWEDY 213

Db 644 LSNKDSITRLYYGDMSDDGQYMATKSPYYDAIDLKARIKYAAGQDMKITVYEGDKS 703

Qy 214 NNGWTWA---QSCYYKRG---DGSESTRWLAAGSIPPPONYDGLWL---DNDTALS 263



CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: Curapatsed formatter Version 0.9  
SEQ ID NO 330  
LENGTH: 1306  
TYPE: PRT  
ORGANISM: *Saccharomyces cerevisiae*  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0).-(0)  
OTHER INFORMATION: Polypeptide Accession Number YGR014W  
US-09-538-092-330

Query Match 5.8%; Score 104; DB 2; Length 1306;  
Best Local Similarity 22.4%; Pred. No. 1.2; Matches 46; Conservatve 43; Mismatches 84; Indels 32; Gaps 9;  
Matches 46; Conservatve 43; Mismatches 84; Indels 32; Gaps 9;

Qy 3 LREDSMSDWGVSFTNVILEGLDNCNTYANGLNMIGVINTPTDDGNNFVFDIDVTLN 62  
Db 269 LQTSSESSSTTAASHALPVNSTDVG-----SSASPVMSAAGQTASSSTDNPMS 320  
Qy 63 DNIKIVD-YDGSDIDG-----SDGWFYTGPNPNEYNTIPNSQSYLSIKSENSQITQIKR 115  
Db 321 ETRSLTSTEVDGSDVSVTVALISAPFLQTSTSNSFIV-SPSVSFVPSQSS----- 372  
Qy 116 YVSCSNTSRIRTKRSFS-AKVTITSGKVISTQN---SINSSRVVINAIDATNFTDEL 169  
Db 373 DVASSTANWVSSSFSDFPPQTSQSGSVWSAQSAHALFOSSTEVIGA-SASSTMSSL 431  
Qy 170 RTT--KETRFENGSYTHKSSTWSL 192  
Db 432 STTSLQSTTLDSSSLASSASSSDL 456

Search completed: January 30, 2006, 09:50:35  
Job time : 8.48879 sec

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
 Copyright (c) 1993 - 2006 Compugen Ltd.

## OM protein - protein search, using SW model

Run on: January 30, 2006, 09:43:19 ; Search time 20.8022 Seconds  
 Sequence: (without alignments) 6708.668 Million cell updates/sec

Title: US-09-889-874a-22

Perfect score: 1784

Sequence: FTLRBDSMSDWYGYSTFENI.....ILFHAGYENSDPYLSSIVY 334

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
 Listing first 45 summaries

Database : Published Applications AA\_Main.\*

1: /cgn2\_6\_ptodata/1/pubpaa/us07\_pubcomb.pep:\*

2: /cgn2\_6\_ptodata/1/pubpaa/us08\_pubcomb.pep:\*

3: /cgn2\_6\_ptodata/1/pubpaa/us09\_pubcomb.pep:\*

4: /cgn2\_6\_ptodata/1/pubpaa/us10\_pubcomb.pep:\*

5: /cgn2\_6\_ptodata/1/pubpaa/us10b\_pubcomb.pep:\*

6: /cgn2\_6\_ptodata/1/pubpaa/us11\_pubcomb.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	126	7.1	892	4 US-10-282-122A-70481 Sequence 70481, A
2	126	7.1	892	4 US-10-282-122A-809-21 Sequence 21, April
3	126	7.1	930	4 US-10-615-383-10 Sequence 10, April
4	126	7.1	930	4 US-10-690-184-10 Sequence 10, April
5	126	7.1	930	4 US-10-699-081-10 Sequence 10, April
6	126	7.1	930	4 US-10-724-972A-5309 Sequence 5309, April
7	124	7.0	1092	4 US-10-378-674-8 Sequence 8, April
8	124	7.0	1092	4 US-10-805-281-15 Sequence 15, April
9	118.5	6.6	670	4 US-10-282-122A-70444 Sequence 70444, April
10	118	6.6	6761	5 US-10-732-923-15035 Sequence 15035, April
11	115.5	6.5	1411	4 US-10-080-505-17 Sequence 17, April
12	115.5	6.5	1411	4 US-10-687-065-17 Sequence 17, April
13	114	6.4	360	5 US-10-617-320-5178 Sequence 5178, April
14	112.5	6.3	1436	4 US-10-080-505-13 Sequence 13, April
15	112.5	6.3	1436	4 US-10-687-065-13 Sequence 13, April
16	111	6.2	369	5 US-10-477-928-1818 Sequence 1818, April
17	109.5	6.1	285	4 US-10-452-024-73 Sequence 108, April
18	109	6.1	807	3 US-09-820-843A-108 Sequence 6661, April
19	108	6.1	626	5 US-10-730-930-6661 Sequence 75356, April
20	108	6.1	978	4 US-10-282-122A-75356 Sequence 418, April
21	107.5	6.0	324	5 US-10-994-726-418 Sequence 417, April
22	107.5	6.0	343	5 US-10-994-726-417 Sequence 73, April
23	107.5	6.0	623	4 US-10-452-024-73 Sequence 292261, April
24	107.5	6.0	657	4 US-10-425-115-292261 Sequence 54358, April
25	107.5	6.0	753	4 US-10-425-114-54358 Sequence 43827, April
26	107.5	6.0	1038	4 US-10-282-122A-75356 Sequence 65711, April
27	107	6.0	1449	4 US-10-282-122A-65711 Sequence 65711, April

## ALIGNMENTS

RESULT 1  
 US-10-282-122A-70481  
 Sequence 70481, Application US/10282122A  
 Publication No. US2004002912A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Karl  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 CURRENT APPLICATION NUMBER: US/10/282.122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIORITY APPLICATION NUMBER: 60/191.078  
 PRIORITY FILING DATE: 2000-03-21  
 PRIORITY APPLICATION NUMBER: 60/206.848  
 PRIORITY FILING DATE: 2000-05-23  
 PRIORITY APPLICATION NUMBER: 60/207.727  
 PRIORITY FILING DATE: 2000-05-26  
 PRIORITY APPLICATION NUMBER: 60/230.335  
 PRIORITY FILING DATE: 2000-09-06  
 PRIORITY APPLICATION NUMBER: 60/230.347  
 PRIORITY FILING DATE: 2000-09-09  
 PRIORITY APPLICATION NUMBER: 60/242.578  
 PRIORITY FILING DATE: 2000-10-23  
 PRIORITY APPLICATION NUMBER: 60/253.625  
 PRIORITY FILING DATE: 2000-11-27  
 PRIORITY APPLICATION NUMBER: 60/257.931  
 PRIORITY FILING DATE: 2000-12-22  
 PRIORITY APPLICATION NUMBER: 60/267.636  
 PRIORITY FILING DATE: 2001-02-09  
 PRIORITY APPLICATION NUMBER: 60/269.308  
 PRIORITY FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 70481  
 LENGTH: 892  
 TYPE: PRT  
 ORGANISM: *Staphylococcus epidermidis*  
 US-10-282-122A-70481

Query Match 7.1%; Score 126; DB 4; Length 892;  
 Best Local Similarity 21.0%; Pred. No. 0..1;  
 Matches 78; Conservative 51; Mismatches 141; Indels 102; Gaps 16;

QY 9 SDWTGUSTVNVLITGLDNCTNLYANGLNMG--VIINITPTDDEGNFVDDVT---- 60  
 Db 398 SMFTNDTNTKHTVEQ---TIYNPLRYSAKETVNLSNGDGGSTI-IDDSIIKVYK 451  
 QY 61 -----LADNIKUVYIDSGDSDGFWYTFGPNEYNT-----IPNS 97  
 Db 452 VGDNONLPDSNRIDYSEVDT-NDDYAOQLGNNNDVNINFNGNIDSPLYIKVSKYDPK 510  
 QY 98 QSYSLKSENSQITOIKRVVSCSNTSRKTSFSAKV--TTSG-----KVIST-- 144  
 Db 511 DDYTIQVTMOTINERY-----TGERRTASYNTIARSTSSQGQGCDLPPEKYKGD 565  
 QY 145 -----TONSINSSRVVINAIDATNPFTDDEBLRTKETRENOVSYSHKSSTNSL-Y 193  
 Db 566 YVWEDVDKOGIONTNDENEKPLSNVLVLTYPDG--TSKSVRTDBEGKQFDGLKNGTY 622  
 QY 194 VHTWTIPSLKLQWRWMDY-----NNGWTAQSCYKTGAGGSSST 236  
 Db 623 KITFETP-----EGYPTPLKHSGTNPALDSEGNVWVTINGQDMTDGSFVQTP 672  
 QY 673 KYSIGNVYWDNTKDGIGQDDEKGISCV--KVTLKDENGNIISTTIDENG--KYQFDN 727  
 Db 297 LDRGHAVIIDQ 308  
 QY 674 KYSIGNVYWDNTKDGIGQDDEKGISCV--KVTLKDENGNIISTTIDENG--KYQFDN 727  
 Db 728 LNSGNVYVHFDFK 739

RESULT 2  
 US-10-661-809-21  
 ; Sequence 21, Application US/10661809  
 ; Publication No. US20040101919A1

GENERAL INFORMATION:  
 ; APPLICANT: HOOK, Magnus  
 ; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR  
 ; TITLE OF INVENTION: POSITIVE BACTERIA  
 ; FILE REFERENCE: P07741US01/BAS  
 ; CURRENT APPLICATION NUMBER: US10/10/661,809  
 ; CURRENT FILING DATE: 2003-09-15  
 ; PRIOR APPLICATION NUMBER: 60/410303  
 ; PRIOR FILING DATE: 2002-09-13  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO: 10  
 ; LENGTH: 892  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis

US-10-661-809-21

Query Match 7.1%; Score 126; DB 4; Length 930;  
 Best Local Similarity 21.0%; Pred. No. 0..11; Mismatches 141; Indels 102; Gaps 16;

QY 9 SDWTGUSTVNVLITGLDNCTNLYANGLNMG--VIINITPTDDEGNFVDDVT---- 60  
 Db 436 SMFTNDTNTKHTVEQ---TIYNPLRYSAKETVNLSNGDGGSTI-IDDSIIKVYK 489  
 QY 61 -----LADNIKUVYIDSGDSDGFWYTFGPNEYNT-----IPNS 97  
 Db 490 VGDNONLPDSNRIDYSEVDT-NDDYAOQLGNNNDVNINFNGNIDSPLYIKVSKYDPK 548  
 QY 98 QSYSLKSENSQITOIKRVVSCSNTSRKTSFSAKV--TTSG-----KVIST-- 144  
 Db 549 DDYTIQVTMOTINERY-----TGERRTASYNTIARSTSSQGQGCDLPPEKYKGD 603  
 QY 145 -----TONSINSSRVVINAIDATNPFTDDEBLRTKETRENOVSYSHKSSTNSL-Y 193  
 Db 604 YVWEDVDKOGIONTNDENEKPLSNVLVLTYPDG--TSKSVRTDBEGKQFDGLKNGTY 660  
 QY 194 VHTWTIPSLKLQWRWMDY-----NNGWTAQSCYKTGAGGSSST 236  
 Db 611 KITFETP-----EGYPTPLKHSGTNPALDSEGNVWVTINGQDMTDGSFVQTP 710  
 QY 237 RWAAGSTPPGNYDGLWDNDIALSGMAHKSTNVDTGQINQLSFRIGKGFSPWVNSG 296  
 Db 711 KYSIGNVYWDNTKDGIGQDDEKGISCV--KVTLKDENGNIISTTIDENG--KYQFDN 765  
 QY 297 LDRGHAVIIDQ 308  
 Db 766 LNSGNVYVHFDFK 777

RESULT 3  
 US-10-615-383-10  
 ; Sequence 10, Application US/10615383  
 ; Publication No. US20040038327A1

GENERAL INFORMATION:  
 ; APPLICANT: FOSTER, Timothy  
 ; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLO  
 ; FILE REFERENCE: P06335T03/BAS  
 ; CURRENT APPLICATION NUMBER: US/10/615, 383  
 ; CURRENT FILING DATE: 2003-07-09  
 ; PRIOR APPLICATION NUMBER: 09/1386, 962  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: 60/098, 443  
 ; PRIOR FILING DATE: 1998-08-31  
 ; PRIOR APPLICATION NUMBER: 60/117, 119  
 ; PRIOR FILING DATE: 1999-01-25  
 ; NUMBER OF SEQ ID NOS: 39  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO: 10  
 ; LENGTH: 930  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis

US-10-615-383-10

Query Match 7.1%; Score 126; DB 4; Length 930;  
 Best Local Similarity 21.0%; Pred. No. 0..11; Mismatches 141; Indels 102; Gaps 16;

QY 9 SDWTGUSTVNVLITGLDNCTNLYANGLNMG--VIINITPTDDEGNFVDDVT---- 60  
 Db 436 SMFTNDTNTKHTVEQ---TIYNPLRYSAKETVNLSNGDGGSTI-IDDSIIKVYK 489  
 QY 61 -----LADNIKUVYIDSGDSDGFWYTFGPNEYNT-----IPNS 97  
 Db 490 VGDNONLPDSNRIDYSEVDT-NDDYAOQLGNNNDVNINFNGNIDSPLYIKVSKYDPK 548  
 QY 98 QSYSLKSENSQITOIKRVVSCSNTSRKTSFSAKV--TTSG-----KVIST-- 144  
 Db 549 DDYTIQVTMOTINERY-----TGERRTASYNTIARSTSSQGQGCDLPPEKYKGD 603  
 QY 145 -----TONSINSSRVVINAIDATNPFTDDEBLRTKETRENOVSYSHKSSTNSL-Y 193  
 Db 604 YVWEDVDKOGIONTNDENEKPLSNVLVLTYPDG--TSKSVRTDBEGKQFDGLKNGTY 660  
 QY 194 VHTWTIPSLKLQWRWMDY-----NNGWTAQSCYKTGAGGSSST 236  
 Db 611 KITFETP-----EGYPTPLKHSGTNPALDSEGNVWVTINGQDMTDGSFVQTP 710  
 QY 237 RWAAGSTPPGNYDGLWDNDIALSGMAHKSTNVDTGQINQLSFRIGKGFSPWVNSG 296  
 Db 711 KYSIGNVYWDNTKDGIGQDDEKGISCV--KVTLKDENGNIISTTIDENG--KYQFDN 765  
 QY 297 LDRGHAVIIDQ 308  
 Db 766 LNSGNVYVHFDFK 777

RESULT 4  
 US-10-690-184-10

```

; Sequence 10, Application US/10690184
; Publication No. US20040141997A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: METHODS FOR TREATING OR PREVENTING INFECTIONS FROM COAGULASE-FILE REFERENCE: P06335US05/BAS
; CURRENT APPLICATION NUMBER: US/10/690,184
; PRIOR APPLICATION NUMBER: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO: 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-10-689-082-10

Query Match 7.1%; Score 126; DB 4; Length 930;
Best Local Similarity 21.0%; Pred. No. 0.11; Mismatches 141; Indels 102; Gaps 16;
Matches 78; Conservative 51; MisMatches 141; Indels 102; Gaps 16;

Qy 9 SDWTGVSTFNVNILETGLDNCNLYANGLMIG--VIINITPTDDEGFDIDVT---- 60
Db 436 SMFTNDTKNHTVEQ---TIVINPLRYSAKETWNISNGDEGSTI-IDSTIIKVYK 489
Qy 61 -----LNDNIKIVDVIDGSIDSGDWYFVNPNYNT-----IPNS 97
Db 490 VEDQNQIPLDSNRVYDVSYEDVT-NDDYAQLGANNNDVINFGNIDSPYIKVSKDPNK 548
Qy 98 QSYSLKSENQSOITOIKYVSCSNTSRRTKFSAKV--TTSG-----KVSI-- 144
Db 549 DDTYTIQQTVMQTINNEY-----TGBFRTASYNTIAFTSSGQGQDLPPEKTYKGD 603
Qy 145 -----TQNSINSSRVVINAIDATNPDELRTKETPENQSYTHGSSTSNSLY 193
Db 604 YWVEDVDKDGIQONTNDKPLSNVNLTVTDG--TSKSVRTDERRGKYODGLKNGLY 660
Qy 194 VHTWTIPSLKLQONWRWEDY-----NNGWTWAQSCYKTCGADGSEST 236
Db 661 KTFETP-----EGYPTPLKHSGTNPALDSEGNSVWTTINGQDMTIDSGFYQTP 710
Qy 237 RMLAAGSIFPPRGNYDGLWDNIALSGMAHKSNVDTGINSFTRIGKGSWVNLISG 296
Db 711 KYSIGNYVWIDTNGDQGDDKEKGISGV--KVTLKDENGNISTTTDENG--KIQFDN 765
Qy 297 LDRGHAVILIDQ 308
Db 766 LNSGNYVWHFDK 777

RESULT 6
US-10-724-972A-5309
Sequence 5309, Application US/10724972A
Publication No. US20040147734A1
GENERAL INFORMATION:
APPLICANT: Doucette-Stamm, Lynn
APPLICANT: Bush, David
TITLE OF INVENTION: NITROIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS FILE REFERENCE: PATH03-16
CURRENT APPLICATION NUMBER: US/10/724,972A
CURRENT FILING DATE: 2003-12-01
PRIOR APPLICATION NUMBER: 09/450,969
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 7544
SEQ ID NO 5309
LENGTH: 930
TYPE: PRT
ORGANISM: S. epidermidis
US-10-724-972A-5309

Query Match 7.1%; Score 126; DB 4; Length 930;
Best Local Similarity 21.0%; Pred. No. 0.11; Mismatches 141; Indels 102; Gaps 16;
Matches 78; Conservative 51; MisMatches 141; Indels 102; Gaps 16;
NUMBER OF SEQ ID NOS: 39
Qy 9 SDWTGVSTFNVNILETGLDNCNLYANGLMIG--VIINITPTDDEGFDIDVT---- 60

```

Db 436 SMFTNIDTKNHTKVEQ----TYINPLRYSAKETVNIVSGNGDEGSTI-IDDSTIKYK 489  
 Qy 61 -----LNDNIKIVDYIDGSDGDGSPWFTSPNPNEYNT-----IPNS 97  
 Db 490 VGNQNLPSDRYDSEVEDT-NDDYAQLGNNNDVNINFNIDSPLYIKVSKYDPNK 548  
 Qy 98 QSYSLKSENSQITQIKRVYCSNTSRRTKPSAKV--TTSG-----KVSI-- 144  
 Db 549 DDTYTIOQVTMOTINIE-----TGBERTASYDNTIAFSTSQQGQGDLPPETKYKGD 603  
 Qy 145 -----TQNSINSSRVINAIDATNPTDDELRTRKETPENQSYTSHGSSTSIL-Y 193  
 Db 604 YWEDVDVKDQGIQONTDNEKPLSNVLVLTYPDG--TSKSVRTDEBKYQFDLKGNLITY 660  
 Qy 194 VHTWTIPRSKLQONWRWEDY-----NNGWTAQSCYKTCRDGSEST 236  
 Db 661 KITPTEP-----EGYTPPLKHSGTNPALDSEGNSVWVTINGQDMTDISGFVQTP 710  
 Qy 237 RMLAAGSTPPPGNYDGMLNDIALSGMAHKSYNVDTGINOLSPTRIGKGFSSWVNISG 296  
 Db 711 KYSLGNYVWYDNTKDGICQGDEKGISGV--KVTLKDENGINIISTTTDENG--KYQFDN 765  
 Qy 297 LDRGHAVITIDQ 308  
 Db 766 LNSGNYIVHFDK 777

RESULT 7  
 US-10-378-674-8  
 ; Sequence 8, Application US/10378674  
 ; Publication No. US2004006209A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PATTI, Joseph M.  
 ; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATIVE  
 ; STAPHYLOCOCCAL PROTEINS  
 ; CURRENT APPLICATION NUMBER: US/10/378,674  
 ; CURRENT FILING DATE: 2003-03-05  
 ; PRIOR APPLICATION NUMBER: 60/361,324  
 ; FILE REFERENCE: P0756US01/BAS  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SEQ ID NO 8  
 ; LENGTH: 1092  
 ; TYPE: PRT  
 ; ORGANISM: *Staphylococcus epidermidis*  
 ; US-10-378-674-8

Query Match Best Local Similarity 7.0%; Score 124; DB 4; Length 1092;  
 Matches 78; Conservative 50; Mismatches 142; Indels 102; Gaps 16;

Qy 9 SDWTGVSTPNVLTGLDNCIYANGLAMIG--VIINITPTDDEGNWVDDVT----- 60  
 Db 439 SMFTNIDTKNHTKVEQ----TYINPLRYSAKETVNIVSGNGDEGSTI-IDDSTIKYK 492  
 Qy 61 -----LNDNIKIVDYIDGSDGDGSPWFTSPNPNEYNT-----IPNS 97  
 Db 493 VGNQNLPSDRYDSEVEDT-NDDYAQLGNNNDVNINFNIDSPLYIKVSKYDPNK 551  
 Qy 98 QSYSLKSENSQITQIKRVYCSNTSRRTKPSAKV--TTSG-----KVSI-- 144  
 Db 604 YWEDVDVKDQGIQONTDNEKPLSNVLVLTYPDG--TSKSVRTDEBKYQFDLKGNLITY 663  
 Qy 194 VHTWTIPRSKLQONWRWEDY-----NNGWTAQSCYKTCRDGSEST 236  
 Db 664 KITPTEP-----EGYTPPLKHSGTNPALDSEGNSVWVTINGQDMTDISGFVQTP 713  
 Qy 237 RMLAAGSTPPPGNYDGMLNDIALSGMAHKSYNVDTGINOLSPTRIGKGFSSWVNISG 296  
 Db 711 KYSLGNYVWYDNTKDGICQGDEKGISGV--KVTLKDENGINIISTTTDENG--KYQFDN 768  
 Qy 297 LDRGHAVITIDQ 308  
 Db 766 LNSGNYIVHFDK 780

RESULT 8  
 US-10-806-288-15  
 ; Sequence 15, Application US/10806288  
 ; Publication No. US200400936A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gubs, Bengt  
 ; APPLICANT: Nilsson, Martin  
 ; APPLICANT: Fryberg, Lars  
 ; APPLICANT: Flock, Jan-Ingvmar  
 ; APPLICANT: Lindberg, Martin  
 ; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from  
 ; TITLE OF INVENTION: Coagulase-Negative *Staphylococcus*  
 ; FILE REFERENCE: GUBS 09/147405  
 ; CURRENT APPLICATION NUMBER: US/10/806,288  
 ; CURRENT FILING DATE: 2004-03-23  
 ; PRIORITY APPLICATION NUMBER: US/09/147,405  
 ; PRIORITY FILING DATE: 1999-04-11  
 ; PRIORITY APPLICATION NUMBER: PCT/SE97/10191  
 ; PRIORITY FILING DATE: 1997-06-18  
 ; PRIORITY APPLICATION NUMBER: SE 9602496-3  
 ; PRIORITY FILING DATE: 1996-06-20  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SEQ ID NO 15  
 ; LENGTH: 1092  
 ; TYPE: PRT  
 ; ORGANISM: *Staphylococcus epidermidis*  
 ; US-10-806-288-15

Query Match Best Local Similarity 21.0%; Score 124; DB 4; Length 1092;  
 Matches 78; Conservative 50; Mismatches 142; Indels 102; Gaps 16;

Qy 9 SDWTGVSTPNVLTGLDNCIYANGLAMIG--VIINITPTDDEGNWVDDVT----- 60  
 Db 439 SMFTNIDTKNHTKVEQ----TYINPLRYSAKETVNIVSGNGDEGSTI-IDDSTIKYK 492  
 Qy 61 -----LNDNIKIVDYIDGSDGDGSPWFTSPNPNEYNT-----IPNS 97  
 Db 493 VGNQNLPSDRYDSEVEDT-NDDYAQLGNNNDVNINFNIDSPLYIKVSKYDPNK 551  
 Qy 98 QSYSLKSENSQITQIKRVYCSNTSRRTKPSAKV--TTSG-----KVSI-- 144  
 Db 604 YWEDVDVKDQGIQONTDNEKPLSNVLVLTYPDG--TSKSVRTDEBKYQFDLKGNLITY 663  
 Qy 194 VHTWTIPRSKLQONWRWEDY-----NNGWTAQSCYKTCRDGSEST 236  
 Db 664 KITPTEP-----EGYTPPLKHSGTNPALDSEGNSVWVTINGQDMTDISGFVQTP 713  
 Qy 237 RMLAAGSTPPPGNYDGMLNDIALSGMAHKSYNVDTGINOLSPTRIGKGFSSWVNISG 296  
 Db 711 KYSLGNYVWYDNTKDGICQGDEKGISGV--KVTLKDENGINIISTTTDENG--KYQFDN 768  
 Qy 297 LDRGHAVITIDQ 308  
 Db 766 LNSGNYIVHFDK 780

RESULT 9  
 US-10-222-122A-7044  
 ; Sequence 7044, Application US/10202122A



```

; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-080-505-17

Query Match 6.5%; Score 115.5; DB 4; Length 1411;
Best Local Similarity 20.8%; Pred. No. 1.5; Mismatches 85; Conservatve 48; MisMatches 131; Indels 145; Gaps 21; Matches 85;

QY 4 REDDSMDWTOVS-----TPNVILETGLDNCNTIANGLANNIGVINIT----- 45
Db 554 RDIAGNGWPGDKDDTKNTGRNLTNP-----NQDGGTLV 607
QY 46 ---PTDBEGNFVD-----IDDVTLANDNIKIVDY-IDG-----SDIDG 78
Db 668 FSGRPFPHAYHNLNLNLGRPKERVVIDDDWINTKAENFQIKGGSTVWSRNVSIEG 667
QY 79 SDGMWFTQNP-BYNTIPNSQSYLSKSENQITOQKIVSNSNTSRKTSFAKVT 137
Db 668 N-WTISMANATFGVUVNQWICTRSDWIGT-----TCKTNLTDKKVIDSIPTQ 719
QY 138 SGKVVISITONS-----INSSRVVIN-----AIDATN- 163
Db 720 INGSINLNATNATNHLGLAKLNGNLINHSQFTLSNNATQHANATVDNANL 779
QY 164 ---FTDBDLRTKETPENOSYTSHKSSTSLSLYVHTWTIPRSKLQWRWEDYNGWT 218
Db 780 NGNVLHFLDSAQFLSKNLSHSHQ-TQGDKDTVTLENATWMPSTDTLQNL---TLLNSTV 835
QY 219 WAQSCY-----YKTGADGGSBSTRW--LAAGSFPPGNYGL---WLDN 257
Db 836 TLNSAYSASSNAPRRRSLETTPTSEBHRNFTLVNGKLSGQGTQFTSSLRGYKSD 895
QY 258 DIAUSGMARKSYN--DTG----INQSFTRIGKGSWWNMSGLD 298
Db 896 KIKLUSNDAEGDYTLAVRDTGKEPVTLQML-----IEGLD 931

RESULT 12
US-10-687-046-17
; Sequence 17, Application US/10687046
; Publication No. US20040157241A1
; GENERAL INFORMATION:
; APPLICANT: St. Game, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59341-1-RFT-BCF/DR
; CURRENT APPLICATION NUMBER: US/10/687, 046
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US/10/080, 505
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296, 791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839, 996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-687-046-17

Query Match 6.5%; Score 115.5; DB 4; Length 1411;
Best Local Similarity 20.8%; Pred. No. 1.5; Mismatches 85; Conservatve 48; MisMatches 131; Indels 145; Gaps 21; Matches 85;

QY 46 ---PTDBEGNFVD-----IDDVTLANDNIKIVDY-IDG-----SDIDG 78
Db 668 FSGRPFPHAYHNLNLNLGRPKERVVIDDDWINTKAENFQIKGGSTVWSRNVSIEG 667
Db 720 INGSINLNATNATNHLGLAKLNGNLINHSQFTLSNNATQHANATVDNANL 779
QY 138 SGKVVISITONS-----INSSRVVIN-----AIDATN- 163
Db 780 NGNVLHFLDSAQFLSKNLSHSHQ-TQGDKDTVTLENATWMPSTDTLQNL---TLLNSTV 835
QY 219 WAQSCY-----YKTGADGGSBSTRW--LAAGSFPPGNYGL---WLDN 257
Db 836 TLNSAYSASSNAPRRRSLETTPTSEBHRNFTLVNGKLSGQGTQFTSSLRGYKSD 895
QY 258 DIAUSGMARKSYN--DTG----INQSFTRIGKGSWWNMSGLD 298
Db 896 KIKLUSNDAEGDYTLAVRDTGKEPVTLQML-----IEGLD 931

```

```

RESULT 13
US-10-617-320-5178
; Sequence 5178, Application US/10617320
; Publication No. US20050136404A1
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02454

COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/617, 320
; FILING DATE: 10-Jul-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107, 433
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/ 085, 311
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneka
; REGISTRATION NUMBER: 40, 485
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEX/FAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae

```

FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (B) LOCATION 1..360  
 SEQUENCE DESCRIPTION: SEQ ID NO: 5178:  
 US-10-617-320-5178

Query Match Best Local Similarity 25.6%; Score 114; DB 5; Length 360;  
 Matches 46; Conservative 33; Mismatches 73; Indels 28; Gaps 8;

Oy 31 YANGLNMGIVIN--ITPDDEGNFDVTDVTDVNDNIKIVDYG----SDIDG 78  
 Db 104 YKNAHQALSVLNLGAIPINENDSVIDEVKVGDNTTISAOVAMVOADLVLTDVG 163  
 Oy 79 SWPFTGPNENTIPNSOSYSLKSENOIQRKVSCN-TSRLRTSAKVT 137  
 Db 164 ---LYTGNP---NSDPRAKRERIETINRETFIDMAGGAGSNGTGMLTKIKATIAE 216  
 Oy 138 SGKVISTONINSNSRVVINAIDATNFIDE--LRTKETTRENOSTHKSSTNSIV 194  
 Db 217 SGPVVIC-SSLKD---SMIEAEETEDGSYFVAQBGKLRQKQMLAFYROSQSIW 271

RESULT 14  
 US-10-080-505-13  
 ; Sequence 13, Application US/10080505  
 ; Publication No. US20030107316A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: St. Geme, Joseph W.  
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS  
 ; INVENTOR: St. Geme, Joseph W.  
 ; FILE REFERENCE: A-5994-1-RFT/DCP/DHR  
 ; CURRENT APPLICATION NUMBER: US/10/080,505  
 ; PRIOR FILING DATE: 2002-02-22  
 ; PRIOR APPLICATION NUMBER: US 08/296,791  
 ; PRIOR FILING DATE: 1994-10-25  
 ; PRIOR APPLICATION NUMBER: US 09/839,996  
 ; PRIOR FILING DATE: 2001-04-20  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO: 13  
 ; LENGTH: 1436  
 ; TYPE: PRT  
 ; ORGANISM: Haemophilus influenzae  
 US-10-687-046-13  

Query Match Best Local Similarity 21.1%; Score 112.5; DB 4; Length 1436;  
 Matches 64; Conservative 39; Mismatches 103; Indels 97; Gaps 14;

Oy 4 REDSMSDTGVSFN-----VILEGLDNQNYIANGLNMGIVINIT----- 45  
 Db 570 KETAYNGWFGGETDENKGNGRLNLYKPTTEDRFLLUSGGTNKG--NITOBGGTLVFSG 626  
 Oy 46 -PTDDEGNFD-----IDDVLTNDNIKIVD-Y-IDG-----SDIDSGD 81  
 Db 627 RPTPHAVHNLNRNELGRPROGEVUVIDDWITRKAENFOIGGSATVRSNSISIEN-- 684  
 Oy 82 WFYTGNPN-EYNTIPNSOSYSLKSENOIQRKVSCN-TSRLRTSAKVT 140  
 Db 685 WTVSNANAAFGVVPNQONTICRSDWGLT-----TCKTVLDLDTKVINSIPTQING 738  
 Oy 141 VISITONS-----INSSRVVIN-----ADATDN---- 163  
 Db 739 SINLTDNATVNTGLAKANGNTLHNHSOFTLSNNTAQTGNTQLSNHANATDNNLNGN 798  
 Oy 164 -PTDDELRITKTRFENQSTYSHKSSTSLSLVHTWTPSLKLQWRWEDYNGWTAQ 221  
 Db 799 VHLTDSAQPSLKNHSFSEQ-1GCDKTVLTLNATWMPSDATLON--TUNNSTVILN 854  
 Oy 627 RPTPHAVHNLRNPLGRPGQEVVDDDWITRKAENFOIGGSAVVSRNTSIEGN-- 684  
 Oy 82 WFYTGNPN-EYNTIPNSOSYSLKSENOIQRKVSCN-TSRLRTSAKVT 140  
 Db 685 WTVSNANAAFGVVPNQONTICRSDWGLT-----TCKTVLDLDTKVINSIPTQING 738  
 Oy 141 VISITONS-----INSSRVVIN-----ADATDN---- 163  
 Db 739 SINLTDNATVNTGLAKANGNTLHNHSOFTLSNNTAQTGNTQLSNHANATDNNLNGN 798  
 Oy 164 -PTDDELRITKTRFENQSTYSHKSSTSLSLVHTWTPSLKLQWRWEDYNGWTAQ 221  
 Db 799 VHLTDSAQPSLKNHSFSEQ-1GCDKTVLTLNATWMPSDATLON--TUNNSTVILN 854  
 Oy 222 SCY 224  
 Db 855 SAY 857

Search completed: January 30, 2006, 09:52:52  
 Job time : 22.8022 secs

RESULT 15  
 US-10-687-046-13  
 ; Sequence 13, Application US/10687046  
 ; Publication No. US20040157241A1  
 ; GENERAL INFORMATION:  
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS  
 ; FILE REFERENCE: 59941-1-RFT/DCP/DHR  
 ; CURRENT APPLICATION NUMBER: US/10/687,046  
 ; CURRENT FILING DATE: 2003-10-15  
 ; PRIOR APPLICATION NUMBER: US/10/080,505  
 ; PRIOR FILING DATE: 2002-02-22  
 ; PRIOR APPLICATION NUMBER: US 08/296,791  
 ; PRIOR FILING DATE: 1994-10-25  
 ; PRIOR APPLICATION NUMBER: US 09/839,996  
 ; PRIOR FILING DATE: 2001-04-20  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO: 13  
 ; LENGTH: 1436  
 ; TYPE: PRT  
 ; ORGANISM: Haemophilus influenzae

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
 Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 30, 2006, 09:46:19 ; Search time 2.66268 Seconds

Sequence: 1 FTLRBDMSDWGTVSTPNTV.....ILFHAGYENSDPYLLSSIVY 334  
 1358.372 Million cell updates/sec

Title: US-09-889-874A-22  
 Perfect score: 1784

Sequence: 1 FTLRBDMSDWGTVSTPNTV.....ILFHAGYENSDPYLLSSIVY 334

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	6.6	773	7 US-11-052-554A-341	sequence 341, App
2	107	6.0	1449	7 US-11-052-554A-237	Sequence 237, App
3	101	5.7	3194	7 US-11-052-554A-90	Sequence 90, Appli
4	100.5	5.6	331	7 US-11-016-564-3	Sequence 3, Appli
5	100.5	5.6	829	6 US-10-880-144-6	Sequence 6, Appli
6	99.5	5.6	2399	7 US-11-052-554-92	Sequence 92, Appli
7	99.5	5.6	2902	7 US-11-052-554A-91	Sequence 91, Appli
8	96.5	5.4	744	7 US-11-052-554A-94	Sequence 94, Appli
9	95.5	5.4	807	7 US-11-052-554A-95	Sequence 55, Appli
10	95.5	5.4	1469	6 US-10-467-657-1088	Sequence 1088, Appli
11	95	5.3	980	7 US-11-052-554A-17	Sequence 17, Appli
12	94	5.3	2314	7 US-11-013-759-11	Sequence 11, Appli
13	93.5	5.2	486	6 US-10-416-047-14	Sequence 14, Appli
14	93.5	5.2	710	7 US-11-089-551A-33	Sequence 33, Appli
15	92.5	5.2	1579	7 US-11-052-554A-9	Sequence 9, Appli
16	92.5	5.2	2053	7 US-11-03-759-9	Sequence 9, Appli
17	91.5	5.1	453	7 US-11-052-554A-224	Sequence 224, Appli
18	91.5	5.1	539	7 US-11-069-642-16	Sequence 16, Appli
19	91.5	5.1	795	7 US-11-052-554A-87	Sequence 87, Appli
20	91.5	5.1	826	6 US-10-793-626-1066	Sequence 1066, Appli
21	2340	5.1	1340	7 US-11-052-554A-171	Sequence 171, Appli
22	90	5.0	395	6 US-10-793-626-668	Sequence 668, Appli
23	90	5.0	1250	7 US-11-052-554A-16	Sequence 16, Appli
24	89	5.0	2715	7 US-11-114-424-51	Sequence 51, Appli
25	88.5	5.0	566	7 US-11-033-039-1244	Sequence 1244, Appli

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/1/pubpea/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpea/US60\_NEW\_PUB.pep:\*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New:\*

1: /cgn2\_6/ptodata/1/pubpea/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/1/pubpea/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpea/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpea/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubpea/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpea/US10\_NEW\_PUB.pep:\*

7: /c

**RESULT 2**  
 US-11-052-554A-237  
; Sequence 237, Application US/11052554A  
; Publication No. US20050288866A1  
**Db**  
**Qy**  
**402** AYERS 406  
; LENGTH: 3194  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori J99  
; US-11-052-554A-90  
; Query Match 5.7%; Score 101; DB 7; Length 3194;  
; Best Local Similarity 18.6%; Pred. No. 4-4; Mismatches 140; Indels 176; Gaps 22;  
; Matches 84; Conservative 51; Mismatches 17;  
**Applicant:** Sachdeva, et al.  
**Title of Invention:** COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE FILE REFERENCE: 30853/40359A  
**Current Application Number:** US/11/052,554A  
**Current Filing Date:** 2005-02-07  
**Prior Application Number:** US 60/589, 227  
**Prior Filing Date:** 2004-07-0  
**Prior Application:** IN 173/DEL/2004  
**Prior Filing Date:** 2004-02-06  
**Number of Seq ID Nos.:** 763  
**Software:** PatentIn version 3.3  
**SEQ ID NO:** 237  
**Length:** 1449  
**Type:** PRT  
**Organism:** Neisseria meningitidis 22491  
; US-11-052-554A-237  
**Query Match** 6.0%; Score 107; DB 7; Length 1449;  
**Best Local Similarity** 21.1%; **Pred.** No. 0-54; **Mismatches** 126; **Indels** 94; **Gaps** 17; **Matches** 71; **Conservative** 46; **N mismatches** 126; **Indels** 94; **Gaps** 17;  
**Qy**  
**9** SDWGTGVSTENVILETGLDNCNLYANGLNMGVINITPDPDESDNFVDIDD-----V 59  
**Db**  
**722** SDWTGILT-NCVEKTTDD-----KVIALSLTKTDISGN-VSLADAHHLNLTGIA 767  
**Qy**  
**60** TLNDNIKIVDYGSDIDGSDGMFWPTGNPNEVNTIPN-SOSYSLIKSENSQTOIKRVS 118  
**Db**  
**768** TNGENLS-----ANGITTYTVSHRNATONGNLSLVGAQAT-FNQATL 808  
**Qy**  
**119** CSNTSRLRTKSFSAKVTTSGKVVISITONS-INSRRVVAIDATNFTDDELTKETRP 177  
**Db**  
**809** NGNTSASGNAWSFNUAQNGLSLTSDNAKANVHSALNG-----NVSLADKAHPF 851  
**Qy**  
**178** ENQSTYSHKSTNSTLKVH---TWTPRSILQLWRWEDYNGNTWAQSCYY--KTGADG 231  
**Db**  
**862** ENSRFTGOLSGSKOTAHLKDSSEWTLPSTGELEGNI---NIDNATITNSAYRHDAAGQT 918  
**Qy**  
**232** GS-----ESTRMLAAGSFPPGMYDGMLNDNTALSGMAHKSYNDGNGINQNSPTRII 284  
**Db**  
**919** GSVDSPTRRRSSRL--SVPTEVESR--NTTUVNGK-----LINGGOTFRM 964  
**Qy**  
**285** GKGFSTM-----VNISGLDRGHAVIILDQ 308  
**Db**  
**965** SELFGYRSOKLAKESSEGTYTTLAVNNTGEPVSLDQ 1001  
**RESULT 3**  
**US-11-052-554A-90**  
**Sequence 90, Application US/11052554A**  
**Publication No. US20050288866A1**  
**GENERAL INFORMATION:**  
**Applicant:** Sachdeva, et al.  
**Title of Invention:** COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE FILE REFERENCE: 30853/40359A  
**Current Application Number:** US/11/052,554A  
**Current Filing Date:** 2005-02-07  
**Prior Application Number:** US 60/589, 227  
**Prior Filing Date:** 2004-07-0  
**Prior Application:** IN 173/DEL/2004  
**Prior Filing Date:** 2004-02-06  
**Number of Seq ID Nos.:** 763  
**Software:** PatentIn version 3.3  
**SEQ ID NO:** 3  
**Length:** 331  
**Type:** PRT  
**Organism:** Candida albicans  
; US-11-016-564-3  
**Query Match** 5.6%; Score 100.5%; DB 7; Length 331;  
**Best Local Similarity** 25.5%; **Pred.** No. 0.3; ;

**Matches** 42; **Conservative** 23; **Mismatches** 55; **Indels** 45; **Gaps** 8;  
**Query** 9 SPMGTVGSTFVNLTETGLCNQNYANGLNNIG--VLIINTPTDEGNFDIDVT---- 60  
**Db** 165 SMFTNIDTKHNTVEQ----TIVINPLRYSAKETVNNTSGNDEGSTI-IDDSITIKYK 218  
**Query** 61 -----LNDNIKVYDYGSSDIDSGDGWFTYGPNPNEYT-----IPNS 97  
**Db** 219 VGINQNQLDSNRVYDVSVDPT-NNDYAQLGNNDVAINFRNIDSPIVIKISKYDPNK 277  
**Query** 98 QSYSLKGSNSQTOIKRVSCTSRLRTKSFAKV--TTSGK 140  
**Db** 278 DYTITIQQTVMQTTINBY---TGEFRTAISNTIAFSTSSGQ 317  
**RESULT 5**  
**US-10-880-144-6**  
**Publication No.** US20050287537A1  
**GENERAL INFORMATION:**  
**APPLICANT:** Svendsen, Allan  
**TITLE OF INVENTION:** Pullulanase variants and methods for preparing such variants with predetermined properties  
**FILE REFERENCE:** 6072-200-US  
**CURRENT APPLICATION NUMBER:** US/10/880,144  
**CURRENT FILING DATE:** 2004-05-28  
**NUMBER OF SEQ ID NOS:** 40  
**SOFTWARE:** PatentIn version 3.3  
**SEQ ID NO** 6  
**LENGTH:** 829  
**TYPE:** PRT  
**ORGANISM:** Bacillus acidopolullulyticus  
**US-10-880-144-6**  
**Query Match** 5.6%; Score 100.5; DB 6; Length 829;  
**Best Local Similarity** 21.5%; Pred. No. 0.92; **Mismatches** 71; **Indels** 71; **Conservative** 44; **Matches** 71; **Score** 100.5; **Pred. No.** 0.92; **Mismatches** 106; **Indels** 109; **Caps** 15;  
**Query** 11 WTCVSTENVILETGDLNCNLYANGLN--ISWVINITPTDGRGNFVDDYTLNDIKI 67  
**Db** 551 WTGT-----SGLSSDOLVTKQOKGKGIGV-----FNDNIR- 582  
**Query** 68 VVVIDGSDIDGSGCWPYTGPNPNTIPN-----SQSYSLKSENQTOIK 114  
**Db** 583 -NGLDGNYVPDKSAQGFGATGDPNQVNTRVNGSISDFTASSETINVTSDNTWIK 641  
**Query** 115 RYVSCSN----TSRLRTKSFSAKVTT-----GKVISTONSISSRVVNAIDA 161  
**Db** 642 --ISASNNDTQADRIRKMDELAQAVVFTSQGPFPMQGGBEMRKTGGNDNS--YNGDS 696  
**Query** 162 TNFTDDELTTRTKEPTRPENOSYSH-----KSTNSTLVLHTWTPRSIK- 204  
**Db** 697 VNGDFDSRRAQPFENVFDDYWSLTHLRDNHPAFRMTTADQIKOMLTFDPSPTWVAFELKN 756  
**Query** 205 -LONWRMED--YNNGTWAGSCYYKIGADGSESTRWLAAGSISPPGNYGLWLNDI 259  
**Db** 757 HAMHDKWNNTIVNTNPNTI-----AQTLLTSPGNTVIGLGNQV 795  
**Query** 260 ALSGMAHKSYNTDTGIMQSFTRFLIGKFS 289  
**Db** 796 GBSKLIGHVNGTVE--VPALS-TILHQOTS 822  
**RESULT 6**  
**US-11-052-554A-92**  
**Publication No.** US2005028866A1  
**GENERAL INFORMATION:**  
**APPLICANT:** Schdeva, et al.  
**TITLE OF INVENTION:** COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE PROTEINS OF THERAPEUTIC POTENTIAL  
**FILE REFERENCE:** 30853/4039A  
**CURRENT APPLICATION NUMBER:** US/11/052,554A  
**PRIOR APPLICATION NUMBER:** US 60/589,227  
**PRIOR FILING DATE:** 2004-07-20  
**PRIOR APPLICATION NUMBER:** IN 173/DEL/2004  
**NUMBER OF SEQ ID NOS:** 763  
**SOFTWARE:** PatentIn version 3.3  
**SEQ ID NO** 91  
**LENGTH:** 2902  
**TYPE:** PRT  
**ORGANISM:** Helicobacter pylori J99  
**US-11-052-554A-91**  
**Query Match** 5.6%; Score 99.5; DB 7; Length 2902;  
**Best Local Similarity** 17.9%; Pred. No. 5.1; **Mismatches** 66; **Indels** 66; **Conservative** 63; **Score** 99.5; **Pred. No.** 5.1; **Mismatches** 93; **Gaps** 14;  
**Query** 15 SPMGVNLTEGDLNCNLYANG--LNMIGVINITPTDDEGNFV-----DI 56  
**Db** 145 ATENGLGASGSNSFTSMPNGTIDVTSGATINVNNSVEGVGRVGSGAGTHGTATLNA 204  
**CURRENT APPLICATION NUMBER:** US/11/052,554A  
**CURRENT FILING DATE:** 2005-02-07  
**PRIOR APPLICATION NUMBER:** US 60/589,227  
**PRIOR FILING DATE:** 2004-07-20  
**PRIOR APPLICATION NUMBER:** IN 173/DEL/2004  
**NUMBER OF SEQ ID NOS:** 763  
**SOFTWARE:** PatentIn version 3.3  
**SEQ ID NO** 92  
**LENGTH:** 2399  
**TYPE:** PRT  
**ORGANISM:** Helicobacter pylori J99  
**US-11-052-554A-92**  
**Query Match** 5.6%; Score 99.5; DB 7; Length 2399;  
**Best Local Similarity** 23.0%; Pred. No. 4.1; **Mismatches** 75; **Indels** 85; **Gaps** 12;  
**Matches** 75; **Conservative** 38; **Mismatches** 128; **Indels** 85; **Gaps** 12;  
**Query** 13 GSVTNPVNLLETGLDNQNYANGLNNIGVINITPTDDEGNFDID----DVTLNMK 66  
**Db** 655 GWNANGNGLTIGNTESVNDNGLWIG-----HGGGRYITGFFSAANIYLNTNPK 704  
**Query** 67 IVDYIDGSDIDES-----DQWY-----TG-NPNEVNTIPNSQSYL 102  
**Db** 705 TEGVNSISDGSGGANITKASDNTIMQDLSKRNAAETVMIQCTASOHSYTTEDATMNV 764  
**Query** 103 LKSENSOITQIKRVSNTSRLRTKSFSRKVTTQSCKVISTQNSINNSRVVNAIDAT 162  
**Db** 765 TUDFSDFNTWIK-----FSFSAKNITSFSNAPSGFTNPQGGSTISTNASNL 811  
**Query** 163 NFDDER-----TKEPTEFENQSY-----TSHKSSTSNLSVHTWIPRSILQ 206  
**Db** 812 SFDSRSLNGGAIYNLOANSLIFTNTOAVENVLVYSGTSNSFNATQNLTSPLSSQIL 871  
**Query** 207 NRWEDDINGWTWAQSCYYKIGADGSESTRWLAAASIFPPENYDGWLDNDIALSGMAH 266  
**Db** 872 NP-----NGDTLQN--NANITLGKSO -AATKNSLTUDNSNSLSDQSVNANGT 920  
**Query** 267 KSYVNDGIN-----QSFTRIGKG 287  
**Db** 921 SAFNNOQASLNTVNGSQMFSSLFPNG 946



RESULT 11  
US-11-052-554A-17  
; Sequence 17, Application US/11052554A  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; FILE REFERENCE: 30853/40359A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US/11/052,554A  
; PRIORITY FILING DATE: 2004-07-20  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2314  
; TYPE: PRT  
; ORGANISM: Moraxella catarrhalis  
US-11-013-759-11  
; Sequence 11, Application US/11013759  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena M.  
; APPLICANT: Sasaki, Ken  
; APPLICANT: Yang, Yan Ping  
; APPLICANT: Klein, Michael H.  
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE  
; FILE REFERENCE: 1038-921MS-1b  
; CURRENT FILING DATE: 2004-12-16  
; PRIOR APPLICATION NUMBER: US/09/361,619  
; PRIOR FILING DATE: 1999-07-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2314  
; TYPE: PRT  
; ORGANISM: Moraxella catarrhalis

Query Match 5.3%; Score 95; DB 7; Length 2314;  
Best Local Similarity 22.0%; Pred. No. 11; Gaps 17;  
Matches 68; Conservative 44; Mismatches 103; Indels 94; Gaps 17;

Query 44 ITPTDDESNFVDDVV---TUNDNIKIVDYIDGSDTDGSDCWFTCNPNEYNTIP--- 95  
DB 748 ITSBDKDKKASQGDINTGPNKNSNSVGTV-----SYNTVUDID 790  
Query 96 -NSQSYSILKSENQSOIQKRVSC-----SNTSLRURTKSPSAKVUTTSKGKIS 143  
DB 791 GNAITAKTYDTEQNQTSKVTYDVNDBEKTIELTGDGNGKINKIGVKTTLTNNANGKA- 848  
Query 144 ITONSINSRVVINALDT--NFTEDELRTKET-----RENOQVSHK 186  
DB 849 -TNFSTTDNDALVNAKDTAENLTLAKEIHTKGKADTAOFPKVKKGATDETIVGK 907  
Query 187 SST-NSLVLYHTWITPRSKLQNRWEDYNGWTAQSCYYKIGADG---GSBSTRMIAA 241  
DB 908 DGTQNGKTVT----LKLKG-----ENGIVTA----TNGDGTTFEGINTQSGKA 949  
Query 242 GSIFPPGNYDGLMLNDNIALSGMAHKSYNVDTGCGNOLSFTRIGKGSFSSWVNISGLGRH 301  
DB 950 GD-STTLNKDGISKNPAS-----NEQIQVAGDGVFKAK-VDKGNS---STGID-GT 995  
Query 302 AVILIDQG 310  
DB 996 SRITKDQIG 1004

RESULT 13  
US-10-415-047-14  
; Sequence 14, Application US/10416047  
; Publication No. US20050266512A1  
; GENERAL INFORMATION:  
; APPLICANT: Buckley  
; TITLE OF INVENTION: Detection of Proteases and Screening for Protease Inhibitors  
; FILE REFERENCE: 61198-353  
; CURRENT APPLICATION NUMBER: US/10/416,047  
; CURRENT FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: 60/247,160  
; PRIOR FILING DATE: 2000-11-07  
; PRIOR APPLICATION NUMBER: PCT/CA01/01561  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 486  
; TYPE: PRT  
; ORGANISM: Aeromonas hydrophila  
US-10-416-047-14

Query 251 D--GLWLNDLIALSGMAHKSY--VDTGINQLSFTRIGK 287  
DB 395 DENGHIVGNEVYQSGLFNVTLGCGSEMDIT'KSSLIDTISNG 437

Query Match 5.2%; Score 93.5; DB 6; Length 486;  
 Best Local Similarity 19.1%; Pred. No. 1.8; Mismatches 41; Indels 131; Gaps 15;  
 Matches 66; Conservative 31; Mismatches 85; Indels 23; Gaps 5;

Qy 8 MSDA--TGUSTFVNILEGLDN----CNYANGLMI-GVINITPTD-DEGNFTDI 56  
 Db 65 MGQWQISGLANGWIMGEGYNGEIKPGTASNTWCYPTNPVTPSBIPLSALDIPDGDEVY 124

Qy 57 DVTFLNDN--IKIVDYID----GSDID--GSDGMFYTNPN---- 89  
 Db 125 QWLVLHDSAMPKPIETSYLAHVGYAWVGNSHQSQVGEIDMDVTRGDGVIRGENGDG 184

Qy 90 -----EYNTIPNSOSYSLKSENQSOITO-----IKRY 116  
 Db 185 YRCGDKDTAKVSNFAYNLDPDFKHGDVIQSDRQLVKTVGAWVANDSDTPQSGDVFTVY 244

Qy 117 VSCNTSRKTSRSKTT-----GKVISITONI---NSRVVINAIDATMFTDEL 169  
 Db 245 DTATNWNSKNTYGYLSBKVTTKNPKWPLVGETBOLSIETIAANDQWSQNGGSTTLSQSQV 304

Qy 170 RTKETR-----FENSYTSH-----KSTNSLJLVH-----TWT- 198  
 Db 305 RPTVPARSKPVPKELYKADISVYEFKADSVTILSFLRNGCONAMYTHPNRPWNH 364

Qy 199 -----T-TRSLKLQNLQRWEDYNGWWTWQS 222  
 Db 365 TVNIGPYKOKASSIRYQWDKRYIPGBVKWWDWNNTIQQNLSTMQN 410

---

RESULT 14  
 US-11-052-554A-9

Query Match 5.2%; Score 92.5; DB 7; Length 1579;  
 Best Local Similarity 18.9%; Pred. No. 9; Mismatches 121; Indels 103; Gaps 18;  
 Matches 68; Conservative 67; Mismatches 121; Indels 103; Gaps 18;

; GENERAL INFORMATION:  
 ; APPLICANT: Lindquist et al.  
 ; TITLE OF INVENTION: ELECTRICAL CONDUCTORS AND DEVICES FROM PRION-LIKE PROTEINS  
 ; FILE REFERENCE: 30554/40025A  
 ; CURRENT APPLICATION NUMBER: US/11/089, 551A  
 ; CURRENT FILING DATE: 2005-03-24  
 ; PRIOR APPLICATION NUMBER: US 60/559, 286  
 ; PRIOR FILING DATE: 2004-03-31  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO: 33  
 ; LENGTH: 710  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae

US-11-089-551A-33

Query Match 5.2%; Score 93.5; DB 7; Length 710;  
 Best Local Similarity 22.8%; Pred. No. 2.8; Mismatches 41; Indels 131; Gaps 15;  
 Matches 41; Conservative 31; Mismatches 85; Indels 23; Gaps 5;

Qy 70 YDGSDDGSGDGFPTGNPHENY----IPNSQSYSLKSENQSOITOQKRVSCSNTS 123  
 Db 6 YNGNSNVGDT----YDRNDTRNNTRNARSWRSVRSDKGRRSSSTSCKGSYRTRAGRSIDNTNSA 61

Qy 124 RIRTWSPSAKVTTSGKVISITONINNSRVVINAIDATFDDLBRLTKETRENQSYT 183  
 Db 62 KHHSKSKSTVWVTS----SDSNSPTYARVSSTIVATS-----STTRTRRNNTS 110

Qy 184 SHKSSTNSLYVHTWTPRSLKQLOUNWRWEDYNGWTAQSCYKTCGADGSEESTRMALAGS 243  
 Db 111 STASSTTDVGNATSAWSAMASNTSSDQATSYTRKSTDNTA--NSKGNNWSSAGN 168

RESULT 15  
 US-11-052-554A-9  
 ; Sequence 9, Application US/11052554A  
 ; Publication No. US2005028866A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sachdeva, et al.

Query Match 5.2%; Score 93.5; DB 6; Length 486;  
 Best Local Similarity 19.1%; Pred. No. 1.8; Mismatches 41; Indels 131; Gaps 15;  
 Matches 66; Conservative 31; Mismatches 85; Indels 23; Gaps 5;

Qy 146 QNSINSSRUVIN-AIDATNFUDELRTRKTRPENQSYTSH-----KSTNSLJLVH----- 198  
 Db 584 NNDVYADCKTENTLTINTIDESNPL-TNHQVTFKNKGSAETVEPPQONTDAYGVATIN 642

Qy 199 IPSLKLQNLQRWEDYNGWWTWQSOCYKTCGADGSEESTRMALAG----- 242  
 Db 643 MVSQAVENTISATLPG--FSQRIARPKVSSTPKPKOLVADPDTIAGNSOGSTLTA 700

Qy 243 -----SIRPPGNYDGLWINDIA-----LSGMAKHSYNDTGI 275  
 Db 701 IITDFHNPNPLKDMKVNFWPG--GSQIDNTATDQSGIVRVHLTSSKAGSYSDASL 756

---

Search completed: January 30, 2006, 09:53:18  
 Job time : 3.66268 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2006, 09:41:43 ; Search time 5.82461 Seconds

Sequence: 1 FTLRBDMSDWGTVSTPNNVIL...ILFHAGYENSDPYLSSSIVY 334  
5517.344 Million cell updates/sec

Title: US-09-889-874A-22  
Perfect score: 1784

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 Seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : PIR 80:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result NO.	Score	Query Match Length	DB ID	Description
1	124	7.0	T30214	fibrinogen-binding protein - <i>Staphylococcus epidermidis</i>
2	122.5	6.9	E89433	CSpecies: <i>Staphylococcus epidermidis</i>
3	117.5	6.6	B32335	CAccesion: T30214
4	115.5	6.5	S71228	Infect. Immun. 66: 2666-2673, 1998
5	115	6.2	R95107	A Title: A Fibrinogen-binding protein of <i>Staphylococcus epidermidis</i> .
6	111	6.2	H97975	A Reference number: 220781; Muid: 98261511; PMID: 9596732
7	109	6.1	B77605	A Status: preliminary, translated from GB/EMBL/DBJ
8	109	6.1	T30857	A-Molecule type: DNA
9	109	6.1	I449	A Residues: 1-102 <NHL>
10	108	6.1	T30552	A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR
11	107.5	6.0	E70149	RESULT 1
12	107.5	6.0	I038	T0214
13	107	6.0	H90053	fibrinogen-binding protein - <i>Staphylococcus epidermidis</i>
14	107	6.0	A86227	CSpecies: <i>Staphylococcus epidermidis</i>
15	107	6.0	JH0284	CAccesion: T30214
16	106	5.9	B81963	Infect. Immun. 66: 2666-2673, 1998
17	105.5	5.9	A32935	A Title: A Fibrinogen-binding protein of <i>Staphylococcus epidermidis</i> .
18	105	5.9	G97336	A Reference number: 220781; Muid: 98261511; PMID: 9596732
19	105	5.9	507	A Accession: T30214
20	105	5.9	A57622	PIR 80:*
21	105	5.9	E69733	1: pir1:*
22	104	5.8	A45866	2: pir2:*
23	104	5.8	D99182	3: pir3:*
24	103.5	5.8	I306	4: pir4:*
25	103.5	5.8	S25370	
26	103.5	5.8	A69088	
27	103	5.8	T39941	
28	103	5.8	A29339	
29	102.5	5.7	H95076	
30			101	102.5
31			102	5.7
32			102	934
33			102	2
34			101.5	5.7
35			101	2894
36			101	2
37			101	5.7
38			101	1136
39			101	1
40			101	US881
41			101	D81019
42			101	5.7
43			100	1649
44			100	2
45			99.5	1577
				2
				2329
				2
				962
				2
				AG2444

#### ALIGNMENTS

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

CCSpecies: *Staphylococcus epidermidis*

CAccesion: T30214 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

RNilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66: 2666-2673, 1998

A Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A Reference number: 220781; Muid: 98261511; PMID: 9596732

A Status: preliminary, translated from GB/EMBL/DBJ

A-Molecule type: DNA

A Residues: 1-102 <NHL>

A Cross-references: UNIPROT:070022; UNIPARC:UPI0000AFC6; EMBL:Y17116; NID:61296734; PIR

Query Match Similarity 7.0%; Score 124; DB 2; Length 1092; Best Local Similarity 21.0%; Pred. No. 0; M88; Mismatches 142; Indels 102; Gaps 16; Matches 78; Conservative

RESULT 1

T0214

fibrinogen-binding protein - *Staphylococcus epidermidis*

C;Species: Sulfolobus solfataricus  
C;Accession: E90433 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awazey, M.J.; Chan-Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, E.;Arrett, R.A.; Ragan, M.A.; Seneca, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A;Description: Sulfolobus solfataricus complete genome.  
A;Reference number: A99139  
A;Accession: E90433  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-990 <KUR>  
A;Cross-references: UNIPROT:Q97WMO; UNIPARC:UPI00000647BD; GB:AB006641; NID:913815906; E;Gene: SSO2602

Query Match 6 9%; Score 122.5; DB 2; Length 990;  
Best Local Similarity 20.8%; Pred. No. 0.66; Indels 99; Gaps 13;  
Matches 64; Conservative 45; Mismatches 99; Indels 99; Gaps 13;

Qy 36 NMIGVINITPTDDEGNGNFDIDD-----VTLNDNIKIVDYGSDIDGSDGW 82  
Db 728 NLNVNLNSLNGKREPSN--NVLIVGPVSLKLVHVLQVYVNLNSNPKALNSTVPSPGW 785  
Qy 83 FTYGNP---NEYNTIPNSQSYSL-----RSSENSQTOIKRYVSCSNTSRLLRK 128  
Db 786 YNAVTPISFINMTYVSNNTYVILSILPNNFTVNRSLTVKTTIKEYL----- 834  
Qy 129 SFSAKVTTISGKVISTONSISSRVVINAIDATNFDDLEURTTKTRFENQSYSHKSS 188  
Db 835 -----VTVNBPPLITANRNTLSTEWPA-----GQTLIIPKVYNSNERFYNTS 882  
Qy 189 TNSLYVHTWITPRSKQLQNWRWEDVINGWTQACSYKTKGADGGSFSTRMLAAGS---- 243  
Db 883 S---VLTINTQOPTSINVK-----PIEYLVLTIDGNB--WLPRGSVTIL 922  
Qy 244 -----FPFGNYDGLWLNDNTIALSGMAHKSYVNDT3GNOQSFTRIGKGFSPWYNNSGLDR 299  
Db 923 QSVPIYEQGKWE-----SYVNSNSV-AITVNQPTETF--VKNINGSFV 964  
Qy 300 GHAVVII 306  
Db 965 GSVILLI 971

RESULT 3  
B33935 hypothetical protein NPL - Entamoeba histolytica (fragment)  
C;Species: Entamoeba histolytica  
C;Accession: B33935 #sequence\_revision 20-Dec-1989 #text\_change 09-Jul-2004  
R;Tannich, E.; Horstmann, R.D.; Knobloch, J.; Arnold, H.H.  
Proc. Natl. Acad. Sci. U.S.A. 86, 5118-5122, 1989  
A;Title: Genomic DNA differences between pathogenic and nonpathogenic Entamoeba histolytic  
A;Reference number: A32935; MNUID:93296555; PMID:2544890  
A;Accession: B33935  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-640 <TRAN>  
A;Cross-references: UNIPROT:P20301; UNIPARC:UPI0000125CB8

Query Match 6.6%; Score 117.5; DB 2; Length 640;  
Best Local Similarity 21.9%; Pred. No. 0.83; Mismatches 73; Conservative 45; Mismatches 121; Indels 95; Gaps 18; Matches 17; Conservatve 73; Mismatches 121; Indels 95; Gaps 18;

Qy 25 LDNCNTYANG--LNMGIVININITPTDDEGNGNFDIDDVTLNDNIKIVDYGSDGW 82  
Db 1347 TAIDLLESSKSTENQDIDV---CSDKSTFVQ-----QILDVYTEKV----I 1387  
Qy 83 FTYGNP---NEYNTIPNSQSYSL-----RSSENSQTOIKRYVSCSNT 122  
Db 1388 LYGVDPNSKVN--ENSKVSYLVTPTHSKLSSILSKNLKNSKFITTNTNT 1445  
Qy 123 SRKRTKPSAKVTTISGKVISTONSISSRVVINAIDATNFDDLEURTTKTRFENQSY 182  
Db 1446 NNIEKNS---SIDSPLSNTSTSSSITPILTSNNLDLN----- 1481  
Qy 183 TSHKSSTSLSVYHTWITPRSLQCLQNWRWEDVINGWTQACSYKTKGADGGSFSTRMLAAG 242  
Db 1482 -NNNNNNNLSLV-----NNG-----GVGGNNVPSTLTTI 1511  
Qy 243 STPPGNYDGLWLNDNTIALSGMAHK---SYNDTGINOLSFTRIG 285  
Db 1512 QOSQPKCYVILAVBDNDNIKVVQRQERKLGYYTAIVGNGLKAEBIG 1558

**RESULT 5**  
**R95107** glutamate 5-kinase [imported] - *Streptococcus pneumoniae* (strain TIGR4)  
**C;Species:** *Streptococcus pneumoniae*  
**C;Accession:** R95107 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
**R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid**  
**on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,**  
**Science 293, 498-505, 2001**  
**A;Authors:** Loftus, B.J.; Yang, P.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
**A;Title:** Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.

**A;Reference number:** A95000; **MUID:**21357209; **PMID:**11463916  
**A;Accession:** F95107  
**A;Status:** preliminary  
**A;Molecule type:** DNA  
**A;Residues:** 1-369 <KUR>  
**A;Cross-references:** UNIPROT:Q97955; UNIPARC:UPI0000051644; GB:AB005672; PIDN:AAK75055.1;  
**A;Experimental source:** strain TIGR4  
**A;Genetic:**  
**C;Superfamily:** glutamate 5-kinase

**Query Match** 6.2%; **Score** 111; **DB** 2; **Length** 369;  
**Best Local Similarity** 25.0%; **Pred.** No. 1 2; **Matches** 45; **Conservative** 34; **Mismatches** 73; **Indels** 28; **Gaps** 8;

<b>Oy</b>	31 YANGLNLMGIVIN-ITPTDDGIGNPFDVPPDVTLLNDNIKIVDGG----- SPIDG 78
<b>Db</b>	113 YKNAHQALSVLNLNGATPI-INDENSDVWIDELKGVDNDTLSAQCVAAMVQADLIVELFDVG 172
<b>Oy</b>	79 SGDWPYFTGNPNEINTINSQSISLKGSENSQTQKRYVCSN-TERRTTSFSAKVTT 137
<b>Db</b>	173 ---LYTGNP--NSDPAKRKRERIETNRIDMAGGAGSNGTSGMLTKIKANTATE 225
<b>Oy</b>	138 SGKVVISITQNSINSSRRVINAIDATNPFDDE---LRTTKEPFENGSYTSHKSTSNSLYV 194
<b>Db</b>	226 SGVFPVYC-SSLKSD---SMIEBAAETEDGSYFVAOEGKLRTOKOMLAFFTAQSOQSIWV 280

**RESULT 6**  
**R97975** glutamate 5-kinase (EC 2.7.2.11) [imported] - *Streptococcus pneumoniae* (strain R6)  
**C;Species:** *Streptococcus pneumoniae*  
**C;Accession:** R97975 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
**R;Hosking, J.A.; Alborn Jr., W.; Arnold, J.; Blazczak, L.; Burgett, S.; Denhoff, B.S.; E**  
**e, R.; LeBlanc, D.J.; Lee, L.N.; LeFKowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M**  
**Y, P.; Sun, P.M.; Winkler, M.E.**  
**J. Bacteriol. 183, 5709-5717, 2001**  
**A;Authors:** Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; A;Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.  
**A;Reference number:** A97872; **MUID:**21429245; **PMID:**11394234  
**A;Accession:** R97975  
**A;Status:** preliminary  
**A;Molecule type:** DNA  
**A;Residues:** 1-376 <KUR>  
**A;Cross-references:** UNIPROT:Q8BQ61; UNIPARC:UPI00000E3527; GB:AE007317; PIDN:AAK99636.1;  
**C;Generic:**  
**A;Gene:** prob  
**A;Superfamily:** glutamate 5-kinase  
**C;Keywords:** phosphotransferase

**Query Match** 6.2%; **Score** 111; **DB** 2; **Length** 376;  
**Best Local Similarity** 25.0%; **Pred.** No. 1 2; **Matches** 45; **Conservative** 34; **Mismatches** 73; **Indels** 28; **Gaps** 8;

<b>Oy</b>	31 YANGLNLMGIVIN-ITPTDDGIGNPFDVPPDVTLLNDNIKIVDGG----- SPIDG 78
<b>Db</b>	120 YKNAHQALSVLNLNGATPI-INDENSDVWIDELKGVDNDTLSAQCVAAMVQADLIVELFDVG 179

C;Genetics: A;Gene: gtfL

Query Match Best Local Similarity 6.1%; Score 109; DB 2; Length 1449; Matches 80; Conservative 45; Mismatches 113; Indels 114; Gaps 22; QY 17 PNVVLTGLDNCN--IYANGLNMGIVIN--TPTDDDEGNF---VDIDDTLNDNTKI 67 Db 459 PEFILLANDIDNSNPRAVEBQNLWLYHIMIGITGGSSEBNPDCVRVDAVNNAIDLQI 518 QY 68 V-DVID--GSD-----LGSDGMYTGPNEYN----TIPNSQ-----YSL 102 Db 519 ASDYFKAKYGAQDOSQDQAQKHLSTLEAHSR--NDAYVNEEDTKGAQLPMDPMPHLALVSL 576 QY 103 LK--SENSQITQI-----KRYVCSNTSRRTKSP---SAKUTTISGKVUSITQNSIN 150 Db 577 LRPIGNRSCVEPLISLNRDSESGKNSKRMANVAFVRAHDSEVOSIQIL---KNEIN 633 QY 202 -----SLLKQWRWEDYDNNGMTWA--Q 221 Db 688 YGDMYTDDGQYMAQSKPSYDAIEFLLKGRIRRAGGQDMKV-NYIGYGTNGMDAAGVLT 746 QY 222 SCYKTKGA---DGCESESTRWLAQGSIFFPGNYQGLMLNDIALS-GMAHKs 268 Db 747 SVRIGTGANSASDCTAETR--NOGMAVIUNQPALRITSNLTNMGAAHRN 796

RESULT 9 T30552 glucosyltransferase N - Streptococcus salivarius (fragment) C;Species: Streptococcus salivarius C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004 C;Accession: T30552 R;Jaffee, R.I. submitted to the EMBL Data Library, February 1998 A;Description: Streptococcus salivarius V1477 gtfN. A;Reference number: Z20854 A;Accession: T30552 A;Status: preliminary; translated from GB/EMBL/DDJB A;Molecule type: DNA A;Residues: 1-626 <VVS> A;Cross-references: UNIPROT:O64599; UNIPARC:UPI00000A2DF3; EMBL:AC03671; NID:g2833627; I C;Keywords: glycosidase; hydrolase; polyaccharide degradation

Query Match Best Local Similarity 6.1%; Score 108; DB 2; Length 626; Matches 65; Conservative 37; Mismatches 122; Indels 90; Gaps 14; QY 26 DNCFNIYANGLNMGIVINNITPTDDDEGNFVDDTLYNNIKTVYDGS----- 74 Db 172 ENSVNUVEEGFSVNSYGT-----GIGAGADTINYQVNWNPFTRSFAYSDGGNGRSQTFKWT 227 QY 75 -DIDGSDPMFYTGPNPNEVNTINS-QSVLSLKSSENSQITQIKYVCSNTSRRTKSSA 132 Db 228 HEANAGNCGSFTSYGKONGCNGVNPBEFTSYGV--SSNVICGGSFSVNGESEN-----AA 276 QY 133 KVTTTS-GKVVISITQNSTNSRNVIN-----DIAINFDTDBLRKTKEFRENQSYT 183 Db 277 NDFTTSYSSDGNVTPQNNTNNGYASGNANVDTFANYRDVANVGDD-----SFS 323 QY 184 SHKESTNSLIVHTWTPRSKLQWNRWHDYDNNGMTWAGSCY--YKTGADGGSSETRWJAA 241 Db 324 SYAKDSNSRSRKVN-----FNYGOSFNPGSFTFGYKGKLAGSKLSFKTTP 369 C;Genetics: gtfN

Query Match Best Local Similarity 6.1%; Score 109; DB 2; Length 1449; Matches 80; Conservative 45; Mismatches 113; Indels 114; Gaps 22; QY 17 FNVVLTGLDNCN--IYANGLNMGIVIN--TPTDDDEGNF---VDIDDTLNDNTKI 67 Db 459 PEFILLANDIDNSNPRAVEBQNLWLYHIMIGITGGSSEBNPDCVRVDAVNNAIDLQI 518 QY 68 V-DVID--GSD-----LGSDGMYTGPNEYN----TIPNSQ-----YSL 102 Db 519 ASDYFKAKYGAQDOSQDQAQKHLSTLEAHSR--NDAYVNEEDTKGAQLPMDPMPHLALVSL 576 QY 103 LK--SENSQITQI-----KRYVCSNTSRRTKSP---SAKUTTISGKVUSITQNSIN 150 Db 577 LRPIGNRSCVEPLISLNRDSESGKNSKRMANVAFVRAHDSEVOSIQIL---KNEIN 633 Db 515 SRRVWNIAIDATNFDDDELRT-----TKEITFENQSYTSHK--SSTNSLIVHTWTPR-- 201 Db 634 PQST-----GNTFTLDEMKAPEIYKNDMSANKQTYQINIPSAYALMULTHKOTPVRY 687 QY 202 -----SLKQWRWEDYDNNGMTWA--Q 221 Db 688 YGDMYTDDGQYMAQSKPSYDAIEFLLKGRIRRAGGQDMKV-NYIGYGTNGMDAAGVLT 746

RESULT 10 T01485 probable polygalacturonase (EC 3.2.1.15) 1 beta chain F1707.9 - Arabidopsis thaliana C;Species: Arabidopsis thaliana (mouse-ear cress) C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 31-Dec-2004 C;Accession: T01485 R;Vysotskai, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li, R.; Li, Y.; Palm, C.J.; Shin, P.; Sun, H.; Davis, R.W.; Becker, J.R.; Federspiel, N. submitted to the EMBL Data Library, June 1998 A;Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence. A;Reference number: Z14334 A;Accession: T01485 A;Map position: 1 A;Intron: 17/3 A;Molecule type: DNA A;Residues: 1-626 <VVS> A;Cross-references: UNIPROT:O64599; UNIPARC:UPI00000A2DF3; EMBL:AC03671; NID:g2833627; I C;Keywords: glycosidase; hydrolase; polyaccharide degradation

Query Match Best Local Similarity 20.7%; Pred. No. 3, 9; Matches 65; Conservative 37; Mismatches 122; Indels 90; Gaps 14; QY 26 DNCFNIYANGLNMGIVINNITPTDDDEGNFVDDTLYNNIKTVYDGS----- 74 Db 172 ENSVNUVEEGFSVNSYGT-----GIGAGADTINYQVNWNPFTRSFAYSDGGNGRSQTFKWT 227 QY 75 -DIDGSDPMFYTGPNPNEVNTINS-QSVLSLKSSENSQITQIKYVCSNTSRRTKSSA 132 Db 228 HEANAGNCGSFTSYGKONGCNGVNPBEFTSYGV--SSNVICGGSFSVNGESEN-----AA 276 QY 133 KVTTTS-GKVVISITQNSTNSRNVIN-----DIAINFDTDBLRKTKEFRENQSYT 183 Db 277 NDFTTSYSSDGNVTPQNNTNNGYASGNANVDTFANYRDVANVGDD-----SFS 323 QY 184 SHKESTNSLIVHTWTPRSKLQWNRWHDYDNNGMTWAGSCY--YKTGADGGSSETRWJAA 241 Db 324 SYAKDSNSRSRKVN-----FNYGOSFNPGSFTFGYKGKLAGSKLSFKTTP 369 QY 242 GS1FPPGNYDGLMLNDIALSGMAHKSVNVTGGINOLSPTRIGKGFS---WV-----YN 293 Db 370 NSTF-----KDVAKKGVFAKYKVST-----TANTVGDGKTVNKWIEPGKFR 413 QY 294 ISGLDRGHAVIID 307 Db 414 ESSIKEGTWIPMPD 427

RESULT 11 E70149 hypothetical protein BB0398 - Lyme disease spirochete C;Species: Borrelia burgdorferi (Lyme disease spirochete) C;Date: 13-Feb-1998 #sequence\_revision 13-Feb-1999 #text\_change 09-Jul-2004 C;Accession: E70149 R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, B.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, B.; Bowman, C.; Garland, S.; Fuji, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Nature 390, 580-586, 1997 A;Authors: Smith, H.O.; Venter, J.C. A;Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*. A;Reference number: A70100; MURID:98005943; PMID:9403685 A;Accession: E70149 A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA  
A;Residues: 1-343 <KLE>  
A;Cross-references: UNIPROT:051359; UNIPARC:UPI0000057476; GB:AB001145; GB:AE000783; NII  
A;Experimental source: strain B31  
C;Superfamily: Borrelia burgdorferi hypothetical protein BB0398

Query Match 6.0%; Score 107.5; DB 2; Length 343;  
Best Local Similarity 18.4%; Pred. No. 1.9;  
Matches 44; Conservative 38; Mismatches 80; Indels 77; Gaps 9;

Ov 14 VSTENVILLETGDLNCNIVANGLNMGIVINITPTDDEGNFVFDIDVTLNDNIVKIVDYG 73  
Db 141 INTYKFLWGMIDNLN-----NEBKGNIFYNALSLED---VQDY--- 177  
Ov 74 SDPDGSDEWPFYTPGNPNEYNTIP-----NSQSTSLLSENSQTO 112  
Db 178 -----DESYFY--KKPLISIPRAHLKIDSRDYFNWVKINTVNPNPEFWVYRNLGDLID 229

Ov 113 IRRVSCSNTSRL----RTKSLSAKVTTSKGVISITQNSNSRVVINAIDATNFDD 167  
Db 230 VKEFVLSENITSKLINRIDRNFFIQSQWDQKGK----SNSNTINSFTTMMIRLGGRK 284

Ov 168 ELRTRKETRFENOSYTSHKSSTNSLYVW-----TIPRSLKQNWRWEDNGNW 21.9  
Db 285 GIOPAKHLADSSDDISYLESGRDHIEHWYFVKRIVPKD-----PEINNGW 335

**RESULT 12**

H90053 hypothetical protein fnb [imported] - *Staphylococcus aureus* (strain N315)  
C;Species: *Staphylococcus aureus*  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C;Accession: H90053  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Gui, L.; Oguri, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: H90053  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1018 <KUR>  
A;Cross-references: UNIPROT:Q99RD2; UNIPARC:UPI00000CAB43; GB:BA000018; PID:g13702453; F  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: fnb

Query Match 6.0%; Score 107.5; DB 2; Length 1072;  
Best Local Similarity 21.2%; Pred. No. 9.7; Matches 65; Conservative 52; Mismatches 136; Indels 54; Gaps 10;

Ov 10 DWGVSTENVILLETGDLNCNIVANGLNMGIVINITPTDDEGNFVFDID 57  
Db 610 DSBRISITSGAFISNSQTSSEASNSNSMSNSINSLSLISLTSATNSNSBTKDN 669  
Ov 59 DVTLDNNDKIVDYGDSIDG----SDGPFYTPGNPNEYNTIPNSQSYILKSENQTOIK 114  
Db 670 SSTHSNNILNSGNSDSSDSDSDSDSDSSNLSSSPN----LETNQTISSKPSEVNNSNP 725  
Ov 115 RVVSCTN-----SRIRKPSKSVKTTSGKVISITO-NSTNSRVVINAIDATN 163  
Db 726 KVYSSNSNVOENSTDHEMSTNPKSISSPISTSSSOKESENQNLNTTEGINNPITTNN 785  
Ov 164 FTDBDLRKTFRENOSYTSHKSSTNSLYVH-----TWIPRSLKQNWRD 212  
Db 766 SSENNSAASLITYSNN-----SSESETGCLYIINEARQDNGSELSHSISSLSSNENTN 838  
Ov 21.3 YANGWTWIAQSCYKRTGAOGSSESTRWLLAGSTPPPGNYDGLWLDNDTALSGMAHKSTND 272  
Db 839 -NISIQQAILESSKSSTNKASSLSIINSTSHPQE---DNQSNSSDEVKSNNNV 892

Ov 273 TGINQLS 279  
Db 893 SILGQLN 899

**RESULT 13**

JH0284 135K surface antigen M17 precursor - *Entamoeba histolytica*  
C;Species: Entamoeba histolytica  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C;Accession: JH0284  
Db 309 NGBKQKITSKINGBEETEKITPVVY--NPGVSNSYTNVNGSIEETNKESENKFHT-AVIRKPM 365  
Ov 121 NTSLRLTKFSKAKVTT-----GKVISITQNSNSRVVINAIDATNF 165  
Db 366 NGNQNTVSVTGUTLGEESNLAGGSOPTVKYTYEGLKDELPOS-----VYANISDTNPK 419

Ov 166 DDELRTKETR-----FENOSYTSHKSSTNSLYV-HWTIPRSLKQNWRWEDYN--- 214  
Db 420 D-----VTKEMNGKLSQLQNGSYISLNLDKDKTVIHYGEVYLOGSDQNFRLYGYPER 475  
Ov 215 -----NGW--TWAQS-CYKYKGADGCSSESTRWLAAAGSFIPPGNYDGLWLNDI--- 259  
Db 476 AKYSYYVVGGRLTWDNGLVLVLYSNKADGNGK-----GOIQ-----DNDFEVKE 520  
Ov 260 -----ALSGMAHKSYNVDGNGOL-----FTRIGKG 287

**RESULT 14**

JH0284 135K surface antigen M17 precursor - *Entamoeba histolytica*  
C;Species: Entamoeba histolytica  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C;Accession: JH0284  
R;Edman, U.; Meretz, M.A.; Rausser, S.; Agabian, N.; Meza, I. J. Exp. Med. 172, 879-888, 1990.  
A;Title: Characterization of an immuno-dominant variable surface antigen from pathogenic *Entamoeba histolytica*.  
A;Reference number: JH0284; MUID:90354789; PMID:1696956  
A;Accession: JH0284  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1114 <BLN>  
A;Cross-references: UNIPROT:Q05352; UNIPARC:UPI000007AD25  
A;Experimental source: strain HM1 IMSS

Query Match 6.0%; Score 107; DB 2; Length 1114;  
Best Local Similarity 21.8%; Pred. No. 10; Matches 77; Conservative 37; Mismatches 125; Indels 114; Gaps 19;

Db 965 SELGYRSDKLUJAESEGTYTLLAWNTGNEPVSLDQ 1001  
Search completed: January 30, 2006, 09:49:39  
Job time : 7.82461 secs

Search completed: January 30, 2006, 09:49:39  
Job time : 7.82461 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2006 Compugen Ltd.

### OM protein - protein search, using SW model

Run on: January 30, 2006, 09:40:48 ; Search time 26.6268 Seconds  
 (without alignments)  
 8849.971 Million cell updates/sec

Title: US-09-889-874A-22  
 Perfect score: 1784  
 Sequence: 1 FTLRKEDSMSDWYGYSTFENI.....ILFHAGYENSDPYLSSIVY 334

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

### Searched:

2166443 BeqB, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : UniProt 05.80:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1749	Q9EV8_XENBV	327	2 Q9EV8_XENBV	Xenorhabdus xenophilus
2	145.5	Q9EV8_XENBV	8.2	2 Q9EV8_XENBV	Xenorhabdus xenophilus
3	136	Q9EV8_XENBV	340	2 Q9EV8_XENBV	Xenorhabdus xenophilus
4	129.5	Q9EV8_XENBV	7.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
5	126	Q9EV8_XENBV	7.3	2 Q9EV8_XENBV	Xenorhabdus xenophilus
6	126	Q9EV8_XENBV	7.1	2 Q9EV8_XENBV	Xenorhabdus xenophilus
7	124.5	Q9EV8_XENBV	7.0	2 Q9EV8_XENBV	Xenorhabdus xenophilus
8	124	Q9EV8_XENBV	7.0	2 Q9EV8_XENBV	Xenorhabdus xenophilus
9	122.5	Q9EV8_XENBV	7.0	2 Q9EV8_XENBV	Xenorhabdus xenophilus
10	122.5	Q9EV8_XENBV	6.9	2 Q9EV8_XENBV	Xenorhabdus xenophilus
11	122.5	Q9EV8_XENBV	6.9	2 Q9EV8_XENBV	Xenorhabdus xenophilus
12	121	Q9EV8_XENBV	6.8	2 Q9EV8_XENBV	Xenorhabdus xenophilus
13	119	Q9EV8_XENBV	6.7	2 Q9EV8_XENBV	Xenorhabdus xenophilus
14	119	Q9EV8_XENBV	6.7	2 Q9EV8_XENBV	Xenorhabdus xenophilus
15	118.5	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
16	118.5	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
17	118	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
18	118	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
19	118	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
20	118	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
21	117.5	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
22	117.5	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
23	117.5	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
24	117	Q9EV8_XENBV	6.6	2 Q9EV8_XENBV	Xenorhabdus xenophilus
25	116.5	Q9EV8_XENBV	6.5	2 Q9EV8_XENBV	Xenorhabdus xenophilus
26	115.5	Q9EV8_XENBV	6.5	2 Q9EV8_XENBV	Xenorhabdus xenophilus
27	115.5	Q9EV8_XENBV	6.5	2 Q9EV8_XENBV	Xenorhabdus xenophilus
28	115.5	Q9EV8_XENBV	6.5	2 Q9EV8_XENBV	Xenorhabdus xenophilus
29	115.5	Q9EV8_XENBV	6.5	2 Q9EV8_XENBV	Xenorhabdus xenophilus
30	115.5	Q9EV8_XENBV	6.5	2 Q9EV8_XENBV	Xenorhabdus xenophilus
31	115.5	Q9EV8_XENBV	6.5	2 Q9EV8_XENBV	Xenorhabdus xenophilus

RESULT 1  
 Q9EV8\_XENBV PRELIMINARY; PRT; 327 AA.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA

Morgan J.A.W., Ousley M., Ellis D., Jarrett P.,

RJ

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR

EMBL; AW29651; CACT19492; 1; -; Genomic DNA.

SEQUENCE

327 AA;

3684 MW;

8417B96EB7060D CRC64;

NCBI\_TAXID=40576;

[1]

RN

NUCLEOTIDE SEQUENCE.

RC

STRAIN=173;

RA



QY 140 -KVISITON--SINSSRVVINA---IDATNFDDDEL---RTKETRFENOSYTHK 186  
 ID QSHRIG STAEQ PRELIMINARY; PRT; 892 AA.  
 AC QSHRIG; 185 NKVIA-TDHWSSPFASGLWNAWLNSLDTISNITMKVLPAYKGLKVGULSEWDTGKV 243  
 QY 187 STSNLSLYVTTWTPRSKLQHQNWTRNTVSMNMFTESDVSTKTIGHWNITSKVVDMGYMF- 246  
 DB 244 SDMASMVFETSALOSAGNIGKWNTRNTVSMNMFTESDVSTKTIGHWNITSKVVDMGYMF- 302  
 QY 247 PGNYDGIL-LDN-DIALSGMMHHSYNTDVGINOLSPFRIIGKPSW-VYNISGLDRGA 302  
 DB 303 -HSYKGEBWGDLSLNWDVSNTVMVMPVSSDRGLSLSNPNTK--WKIKNANTLER-- 355  
 QY 303 VVLDQYGNKRYLIFHAGEN 323  
 DB 356 MFSFDRSISFYRVW--ASNEN 374

**RESULT 5**

QSHRIG\_6\_STAEQ  
 ID QSHRIG\_6\_STAEQ PRELIMINARY; PRT; 892 AA.  
 AC QSHRIG\_6;  
 DT 10-MAY-2005 (Tremblrel. 30, Created)  
 DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)  
 DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)  
 DB SarC protein.  
 GN Orderredlocusnames-SEREP0207;  
 OS Bacterium, Firmicutes, Bacillales, Staphylococcaceae, Staphylococcus, NCBI\_TaxID=176279;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX PubMed=15774886; DOI=10.1128/JB.187.7.2426-2438.2005;  
 RA Gill S.R., Fouts D.E., Archer G.L., Mongodin E.P., DeBoy R.T., Revel J., Paulsen I.T., Kolonay J.F., Brinkac L.M., Beanan M.J., Dodson R.J., Daugherty S.C., Madupu R., Angiuoli S.V., Durkin A.S., Haff D.H., Yamamoto J.J., Khouri H., Utterback T.R., Lee C., Dimitrov G., Jiang L., Qin H., Weidman J., Tran K., Kang K.H., Hance I.R., Nelson K.E., Fraser C.M., RT "Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *S. epidermidis* strain.", J. Bacteriol. 187:2426-2438(2005).  
 DR EMBL; CP00029; AAW53571.1; -; Genomic\_DNA.  
 DR TIGR; SER0207; -.  
 DR GO; GO:0009886; C:cell surface; IEA.  
 DR InterPro; IPR008454; Cna\_B.  
 DR InterPro; IPR001859; Gram\_pos\_anchor.  
 DR Pfam; PP0746; Gram\_pos\_anchor; 1.  
 DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 DR KW Cell wall; Complete\_proteome.  
 SQ SSEQUENCE 892 AA; 98746 MW; EC35C932B5BF497 CRC64;

Query Match 7.1%; Score 126; DB 2; Length 892; Best Local Similarity 21.0%; Pred. No. 2.5; Matches 78; Conservative 51; Mismatches 141; Indels 102; Gaps 16;

QY 9 SDMTGVSTENVILTEGLDNCNTYANGLMIG--VIINTPTDDEGNFDIDVT----- 60  
 DB 398 SMFTNDTQNTHVEQ----TIVNPLRYSAKETVNNTSGNGBGSTI-IDPSTIKYK 451  
 QY 61 -----LNDNIKVDYDGSDDGSGDWYTYGPNEYIT-----IPNS 97  
 DB 452 VGDNQNLPSRNKDYSEVEDVT-NDDYAQLGNNNDVNINFGNIDSPVYIKVSKYDPN 510  
 QY 98 QSYLILKSENSTQIQKRVCSNTSLRKTSAKV--TTSG-----KVIST-- 144  
 DB 511 DYYTIQQTWMTTNEY----TGFERTASYNTIAPSTSQQGDLPPBTCKYKGD 565

**RESULT 6**

QKII13\_STAEP  
 ID QKII13\_STAEP PRELIMINARY; PRT; 931 AA.  
 AC QKII13;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Fibringen-binding protein SarG.  
 GN Name=SarG;  
 OS Staphylococcus epidermidis.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcaceae, Staphylococcus, NCBI\_TaxID=1282;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=R28;  
 RA MEDLINE=20340957; PubMed=10878118;  
 RA McCrea K.W., Hartford O., Davis S., Ni Eichin D., Lina G., Speziale P., Foster T.J., Rock M., "The Sarine-aspartate repeat (sar) protein family in *Staphylococcus epidermidis*," Microbiology 146:1535-1546(2000).  
 DR EMBL; AR25042; NAFT501.0; -; Genomic\_DNA.  
 DR PDB; 1R17; X-ray; A/B/C/D=74-98.  
 DR PDB; 1R19; X-ray; A/B/C/D=74-98.  
 DR GO; GO:0009886; C:cell surface; IEA.  
 DR GO; GO:000518; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008454; Cna\_B.  
 DR InterPro; IPR005777; Gpos\_YSIRK.  
 DR InterPro; IPR001859; Gram\_pos\_anchor.  
 DR Pfam; PP0746; Gram\_pos\_anchor; 1.  
 DR Pfam; PP04650; YSIRK\_signal; 1.  
 DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; YSIRK\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 DR KW Cell wall; Complete\_proteome.  
 SQ SEQUENCE 931 AA; 102955 MW; 591E657D97027116 CRC64;

Query Match 7.1%; Score 126; DB 2; Length 931; Best Local Similarity 21.0%; Pred. No. 2.7; Matches 78; Conservative 51; Mismatches 141; Indels 102; Gaps 16;

QY 9 SDMTGVSTENVILTEGLDNCNTYANGLMIG--VIINTPTDDEGNFDIDVT----- 60  
 DB 437 SMFTNDTQNTHVEQ----TIVNPLRYSAKETVNNTSGNGBGSTI-IDPSTIKYK 490  
 QY 61 -----LNDNIKVDYDGSDDGSGDWYTYGPNEYIT-----IPNS 97  
 DB 491 VGDNQNLPSRNKDYSEVEDVT-NDDYAQLGNNNDVNINFGNIDSPVYIKVSKYDPN 549  
 QY 98 QSYLILKSENSTQIQKRVCSNTSLRKTSAKV--TTSG-----KVIST-- 144  
 DB 550 DYYTIQQTWMTTNEY----TGFERTASYNTIAPSTSQQGDLPPBTCKYKGD 604

QY 145 -----TONSINSSRVVINAIDATNFDDDELRTKETRFENOSYTHK 193  
 DB 566 YVWDVDKDGIONTNWNEKPLSNVLVLTIPDG---TSKSVRTDEEGKQFQDLKNGUY 622  
 DB 194 VHTWTIPSILKLNWRWDY-----NNGWTAOSCYKTGADGGSESTRWLAJAGSFP 236  
 DB 623 KTFETP-----EGYPTPLKHSGTNPALDSEGNSVWVTINGQDDMTIDSFGVCP 672  
 QY 237 RMLAAGSFPPGMYDGLWLDNDLALSMAHKSTVNDTGQINQLSPTRIGKGSWVWMSG 296  
 DB 673 KYSLGNVWYDNTKDGQGDBDKGISY--KVTLKDENGIIISTTUDENG---KYQPDN 727  
 QY 297 LDRCHAVVTDQ 308  
 DB 728 LNSGNYVHFDK 739



QY 145 -----TQNSINSSRVVINAIDATNFTDDEBLRTKETRERENQSTSHSKSTNSL-Y 193  
 Db 607 YWVEDVDKGIGONTNDNEKPLSNLVLTYPDG--TSKSVRTEDECKYQPGKLNGLY 663  
 QY 194 VHTWTIPRSKLQLQWRWEDY-----NNGWTAQSCYTKTGADGSEST 236  
 Db 664 KIRFETP-----EGYPTPLKHSGTNPALDSEGSNVWTINGQDMTIDSGFWQTP 713  
 QY 237 RWLJAGSIFPPGNTDGLWDNDIALSGMAHKSYNVDGNGNQSFTRIGKGSSWVNISG 296  
 Db 714 KYSLGNVWYDTNKDGICGDEDKISGV--KVTLKDENGNIISTTTIDENG--KYQDN 768  
 QY 297 LDGHAVIILIDQ 308  
 Db 769 LNSGNYIVHFDK 780

**RESULT 9**

070022\_STAE  
 ID 070022\_STAE PRELIMINARY; PRT; 1092 AA.  
 AC 070022;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Fibringen-binding protein precursor.  
 DB *Staphylococcus epidermidis*.  
 OS *Staphylococcus epidermidis*.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcaceae; *Staphylococcus*.  
 OX NEBI\_TAXID=1282;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=HB;  
 RX MEDLINE=98261511; PubMed=9596732;  
 RA Nilsson M., Frykberg L., Röck J.I., Pei L., Lindberg M., Guss B.;  
 RT "A fibrinogen-binding protein of *Staphylococcus epidermidis*."  
 RL Infect. Immun. 66:2666-2673(1998).  
 DR EMBL: YI7116; CAA76638.1; -; Genomic\_DNA.  
 DR PIR: T30214; T3014.  
 DR RSPB; Q53653; IN67.  
 DR GO: GO-000986; C:cell surface; IEA.  
 DR GO: GO-0005118; C:cell wall; IEA.  
 DR GO: GO-0016020; C:membrane; IEA.  
 DR InterPro; IPR008454; Cna\_B.  
 DR InterPro; IPR005877; Gpos\_YSLRK.  
 DR InterPro; IPR001699; Gram\_pos\_anchor.  
 DR Pfam; PF05728; Cna\_B; 2.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF04650; YSLRK\_anchor; 1.  
 DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; YSLRK\_anchor; 1.  
 DR PROSITE; PS0847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall; Signal.  
 FT SIGNAL 1  
 FT CHAIN 52 1092 Potential.  
 SQ SEQUENCE 1092 AA; 119292 MW; 654BC39AAB8B984 CRC64;  
 Query Match 6.9%; Score 122.5; DB 2; Length 404;  
 Best Local Similarity 21.1%; Pred. No. 1,7; Gaps 20;  
 Matches 85; Conservative 59; Mismatches 127; Indels 131; Gaps 20;  
 QY 14 VSTENVILLETGLONCNYTANGLMIGVINTAPTDGGNFUDIDDVLTNDNTKIVDYG 73  
 DB 24 IKDFSL-SGSDRQFISNGNQIEFFVN-----VLDDEGEIEEWIKRKDK 72  
 QY 74 SQTGDSQDWPFY--TGMPENEYNT-----PNSQSTSLLKSR----N 107  
 DB 73 KDIQ-ELCWLYRSPGLGOSDEYYSIHGKVALPNAPFKT"QGNPPAQSERLRSKRADQIQYN 131  
 QY 108 SQTQIKY-----USCSNTSRL---RTKSFSAKVTTSGK/-ISITONSINS 151  
 DB 132 SKALQSFVNEYSLPLTWHSQTLSSDWQPTDICAKTSGVITDCYSEPTIV 191  
 QY 1512 SRVVINADATNFTDDEBLRTKETRERENQSTSHSKSTNSL-YVHTWTIPRSKL 205  
 DB 192 SSITPSVIPSQFSSDLDADDIEEVNLORFSSLKALAPALEGFIAYIDSPEE 251  
 QY 206 QNRWEDDNGWTAQSCYTKTGADGSESTRWLASSFPPGNYDGWILWLDN--IALS 262  
 DB 252 PLIIVDEYKVNFACEKTYA-----WIDNEYHFLP 283  
 QY 263 GMHKSY-----NNGWTAQSCYTKTGADGSEST 296  
 DB 284 NLSNPKTITLGKLELNKSVSSKINT-----RGFTHLANVUTPFPYNNMGSYDVTIW 337  
 QY 297 -LDGHAVIILIDQ 308  
 DB 769 LNSGNYIVHFDK 780

**RESULT 10**

Q31IJ9\_9VIBR  
 ID Q31IJ9\_9VIBR PRELIMINARY; PRT; 404 AA.  
 AC Q31IJ9;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Vibrio pomerenensis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.  
 OX NEBI\_TAXID=161725;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=HB-291;  
 RX PubMed=14680703; DOI=10.1016/S0378-1097(03)00843-7;  
 RA Jones J., Apel B., Lewin A.;  
 RT Cloning and molecular characterization of a unique hemolysin gene of *Vibrio pomerenensis* sp. nov.: development of a DNA probe for the detection of the hemolysin gene and its use in identification of *Vibrio* spp. from the Baltic Sea.;  
 RL FEMS Microbiol. Lett. 229:223-229(2003).  
 DR EMBL: AU314791; CAC40977.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 404 AA; 45760 MW; 0EB6B5219A4AC331 CRC64;  
 Query Match 6.9%; Score 122.5; DB 2; Length 404;  
 Best Local Similarity 21.1%; Pred. No. 1,7; Gaps 20;  
 Matches 85; Conservative 59; Mismatches 127; Indels 131; Gaps 20;  
 QY 14 VSTENVILLETGLONCNYTANGLMIGVINTAPTDGGNFUDIDDVLTNDNTKIVDYG 73  
 DB 24 IKDFSL-SGSDRQFISNGNQIEFFVN-----VLDDEGEIEEWIKRKDK 72  
 QY 74 SQTGDSQDWPFY--TGMPENEYNT-----PNSQSTSLLKSR----N 107  
 DB 73 KDIQ-ELCWLYRSPGLGOSDEYYSIHGKVALPNAPFKT"QGNPPAQSERLRSKRADQIQYN 131  
 QY 108 SQTQIKY-----USCSNTSRL---RTKSFSAKVTTSGK/-ISITONSINS 151  
 DB 132 SKALQSFVNEYSLPLTWHSQTLSSDWQPTDICAKTSGVITDCYSEPTIV 191  
 QY 1512 SRVVINADATNFTDDEBLRTKETRERENQSTSHSKSTNSL-YVHTWTIPRSKL 205  
 DB 192 SSITPSVIPSQFSSDLDADDIEEVNLORFSSLKALAPALEGFIAYIDSPEE 251  
 QY 206 QNRWEDDNGWTAQSCYTKTGADGSESTRWLASSFPPGNYDGWILWLDN--IALS 262  
 DB 252 PLIIVDEYKVNFACEKTYA-----WIDNEYHFLP 283  
 QY 263 GMHKSY-----NNGWTAQSCYTKTGADGSEST 296  
 DB 284 NLSNPKTITLGKLELNKSVSSKINT-----RGFTHLANVUTPFPYNNMGSYDVTIW 337  
 QY 297 -LDGHAVIILIDQ 308  
 DB 769 LNSGNYIVHFDK 780

**RESULT 11**  
**Q97VMO\_SULSO** PRELIMINARY; PRT; 990 AA.  
 AC Q97VMO;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Hypothetical protein  
 GN OrderedLocusName=SSO2602;  
 OS Sulfolobus solfataricus  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus;  
 NCBI\_TaxID=2287;  
 RN [1]  
 RP NUCLEOTIDE\_SEQUENCE.  
 RC STRAINATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE:2133296; PubMed:1427726; DOI:10.1073/pnas.141222098;  
 RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
 RA Awazy M.J., Chan Weiher C.C.-., Clausen I.G., Curtis B.A.,  
 RA De Moors A., Brauso G., Fletcher C., Gordon P.M.R.,  
 RA Heikamp-de Jong I., Jeffries A.C., Koera C.J., Medina N., Peng X.,  
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).  
 DR EMBL; AR006856; Ank42724\_1; -; Genomic\_DNA.  
 DR PIR; E90433; E90433; Complete\_proteome; Hypothetical\_protein.  
 KW SEQUENCE 990 AA; 111429 MW; DBE2517A7FBA596A CRC64;

Query Match 6.9%; Score 122.5; DB 2; Length 990;  
 Best Local Similarity 20.8%; Pred. No. 5.1;  
 Matches 64; Conservative 45; Mismatches 99; Indels 99; Gaps 13;

QY 36 NMIGVINTPTDDGDNFVDDD-----VTNDNIKTVVDYTDGSDTDGSDCW 82  
 DB 728 NLVNISLNGKETSN--NVIINGPVLKLYVQYINLNIPKALINSTLTVPSGW 785

QY 83 FYTGSR-----NEYNTIPNQSOSYSLI-----KSENQITOIKRYVSCSNLSRLRK 128  
 DB 786 YNAVTPISFIMPIMTYVNNTRTYILSILPFNFTVNRSLTVKVTILKEYL----- 834

QY 129 SFSAKTTTSGGKVVISITQNSINSRVINADATMFTDELRITKETRKENQSYTHKES 188  
 DB 835 -----VTNPERPLITNTRNTSIVWRP-----GOTLIPKVNLSNERIFYNS 882

QY 189 TNSLVVHTWTIPRSULKQNWRWEDYNGWTAQOSCYYKTGAGGSBSTRMLAAGS---- 243  
 DB 883 S--YLINITQPTPSINSVK-----PIEYLVTDGNE--WLPRPSVITL 922

QY 244 ---IFPPGNIDGLMDNDIALSGMAHKSYNTDGTINQLSFRTRIGKGFSWWNISGLR 299  
 DB 923 QSVPIYEQGKMEG-----SIVNSGV-AITWNQPIETF--VKINGSFV 964

QY 300 GHAVILL 366  
 DB 965 GSIVILL 971

RESULT 12  
**Q7VSS\_PROMM** PRELIMINARY; PRT; 1765 AA.  
 AC Q7VSS;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Hemolysin-type calcium-binding region:RDX N-terminal domain.  
 GN OrderedLocusName=PM0256;  
 OS Prochlorococcus marinus (strain MIT 9313).

OC prochlorococcus;  
 OC NCBI\_TaxID=74547;

**RESULT 13**  
**Q52984\_ECOLI** PRELIMINARY; PRT; 814 AA.  
 AC Q52984;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein precursor.  
 OS Escherichia coli  
 OC Enterobacteriaceae; Escherichia.

OX NCBITaxId=552;

[1] NUCLEOTIDE SEQUENCE.

RP MEDLINE-99412288; PubMed=10482530; RA GO; GO:0016200; C:membrane; IEA.

RA Dutby T.G., Staender L.H., Manning P.A.; RT "55 Silius Biosynthesis Genes from Enterotoxigenic Escherichia coli 115:H40." J. Bacteriol. 181:5847-5851 (1999).

RA EMBL; AU224079; CAA11822.1; -; Genomic\_DNA.

DR GO; GO:0016200; C:membrane; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:006810; P:transport; IEA.

FT HYPothetical protein; Signal.

FT SIGNAL 1 21 Potential.

FT CHAIN 1 1 hypothetical protein.

FT CHAIN 22 814 AA; 90322 MW; C2108F2461B6DB92 CRC64; SQ SEQUENCE 814 AA; 90322 MW; C2108F2461B6DB92 CRC64;

Query Match 6.7%; Score 119; DB 2; Length 814; Best Local Similarity 20.0%; Pred. No. 7.2; Mismatches 94; Conservative 62; Indels 172; Gaps 22; Matches 94;

Qy 1 FTREDSDMSDWGTVSTENY-----TLEGLDNCNTVIA-----NG 34 Db 267 FGTAELBLVDRSYGKTFSTVSPSTGIVEVSKDNLIVLWATQAGYGEINLNALIPLYGQINA 326

Qy 35 LNMT-----GVINITPDE-----GNFVUDDVTLNDNIKV----- 69 Db 327 LVQYKRSQSGIVVVSQNQVNLTTSFSNSDFSWHLEFGNQSSDNFPRKNTTEVIESGVQLP 386

Qy 70 -----YTDGSDIDGSDGWYFTGPNPNEVNTI-----ENSQSLIKKENSQITQ 112 Db 387 VNTLITALYVGGAKDT-----KNTIYSTGLMFQKEPIVSILKGGQQGRHIE 433

Qy 113 IKRVSCSNTSRLRT-----KSFSAKVTTSKGVISITON-SINSRVVIINAT 162 Db 434 MKSVLRLISLWSKKCSTKGWNGKIKSSTONTTISAGYENVNMSVVSANGYIYSSMKEPD 493

Qy 163 NF---TDDELRTKTRPENOSYTHK-----SSTMSLYVHTWTIPRSLKQLQNMW 211 Db 494 YPVANTDHLMESFRYKKNYNSKNLVAMMYNFPGNSLYNTY-----KELRG 545

Qy 212 DY-----NGWTAQSYKTYKTGADGESSEST-----RMLAASSIFPPGY 250 Db 546 DYSVSLGMNISLGKNSRF--NSSFFYKNGADITNSTYAKLSDNWHSVY----- 596

Qy 251 DGIWLDND---IAISGMHK-----YNTDGINOLSF---TRIGK----GP 288 Db 597 -GRFFNDYNSATYTSLSNSNEVRGAGYTYATGQSQULSTDQTLNDSLINSNGIYPTSS 655

Qy 289 SWVYN--ISGLDRQGHAVIILDQYGNKRILFHAGYENSDPYLISSIV 334 Db 656 SWKDQAFIRGKDAKYDISVRNMHDNTIR-----YFSDTNLISPVY 698

RESULT 14

O832P7 ENTPE ID Q832P7 ENTPE PRELIMINARY; PRT; 893 AA..

AC 0832P7; DT 01-JUN-2003 (Tremblel. 24, Created)

DT 01-JUN-2003 (Tremblel. 24, Last sequence update)

DE 01-MAR-2004 (Tremblel. 26, Last annotation update)

DE Conserved domain Protein.

DE Orderable domain names:EF214;

OC Enterococcus faecalis (*Streptococcus faecalis*).

OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.

OX NEBI\_TaxID=1351;

[1] NUCLEOTIDE SEQUENCE.

RP STRAIN=583 / ATCC 700802; MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613; RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seehadri R., Read T.D., Routs D.B., Eisen J.A., Gill S.R., Heidelberg J.F., RA Lee J.C., Hirshman K.; RT whole genome sequencing of *Staphylococcus haemolyticus* uncovers

RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J., RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R., RA Nelson W.C., Vamathevan J.J., Tran J., Hansen T., Upton J., Shetty J., RA Kuo H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A., RA Fraser C.M.; RT "Role of mobile DNA in the evolution of vancomycin-resistant Enterococcus faecalis"; RT Science 299(5607):2071-2074 (2003).

DR EBO16953; AA081906.1; -; Genomic\_DNA.

DR TIGR; EF2174; -.

DR GO; GO:0001796; F:lysozyme activity; IEA.

DR GO; GO:0016998; P:cell wall catabolism; IEA.

DR GO; GO:000953; P:peptidoglycan catabolism; IEA.

DR InterPro; IPR00053; Glyco\_hydro\_25.

DR InterPro; IPR010916; TONB\_Box\_N.

DR Pfam; PF0183; Glyco\_hydro\_25; 1.

DR PRODOM; PD004620; Glyco\_hydro\_25; 1.

DR PROSITE; PS00430; TONA\_DEPENDENT\_REC\_1; UNKNOWN\_1.

DR KW Complete proteome.

SQ SEQUENCE 893 AA; 98579 MW; BCE7D9764AB80DDB CRC64;

Query Match 6.7%; Score 119; DB 2; Length 893; Best Local Similarity 19.4%; Pred. No. 8; Mismatches 74; Conservative 58; Indels 119; Gaps 20; Matches 74;

Qy 21 LEGLDLDN-----CNITYANGLAMIGVINTTPDDEGCFNVD-----DVTIANDNIKV----- 69 Db 138 MNKLGSRRVARATVANVYANDPNLG-----KRNFDIVSSWNGDISHAEVQKTKS 186

Qy 70 Y-IDGSDIDGSDGWYFTGPNPNEVNTI-----ENSQSLIKKENSQITQIKRYVSCNTSRLRT 128 Db 187 YGTVGVSYKLTCTMWV-NPYMAQOIRVAKAGLKVS---ATHYSMTVSA-AQDEAR 240

Qy 129 SFSAKVTTSGKVVISITON-SINSRVVIINAT 162 Db 241 YF-AQAANSG-----LIDKNTIMDNDAPPTLNGRVAHANSVAFNQDILKALGYK 230

Qy 189 TNSLYVHWTI-----PRSIKLOQWRWEDTYNGWTAQSCY 225 Db 291 NDALYVKGWLNTGYYIDTSAFGRDRVWVAQYPVTPDS---SMQWNNDHGAWNQSSQYF 346

Qy 226 KTGDGGSBSTRMLAAGSIFPPG---NYDGLWMDNDIALSGMH-----KSY----- 269 Db 347 -----PGLANTEGPRPDISMVTSNPLANGNSGSPDLSKYTNP 385

Qy 270 -----NYDTGINOLSF-----TRIGKGSWVNVNISGLDRGHAVILIDQ---YG 310 Db 386 GRVIMKNDTTFYQDVAFPTGPGRVVKNTLVTKGIE-YSSAGIPT---LVTDQGUTA 439

Qy 311 NKTRILFHAGYENSDPYISSL 331 Db 440 NKDQVVL--AAQSNIDLYFTN 458

RESULT 15

O4155 STAHJ ID Q4155 STAHJ PRELIMINARY; PRT; 1361 AA..

AC Q4155; DT 13-SEP-2005 (Tremblel. 31, Created)

DT 13-SEP-2005 (Tremblel. 31, Last sequence update)

DT 11-SEP-2005 (Tremblel. 31, Last annotation update)

DB Autolysis.

GN Name=atl; ORName=SH1911;

OS Staphylococcus haemolyticus (strain JCSC1435).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NEBI\_TaxID=279808;

RN [1] NUCLEOTIDE SEQUENCE.

RP STRAIN=JCSC1435;

RA Takeuchi F., Watanabe S., Baba T., Yuzawa H., Ito T., Cui L., RA Morimoto Y., Kurada M., Takahashi M., Anai A., Baba S., Fukui S., RA Lee J.C., Hirshman K.; RT whole genome sequencing of *Staphylococcus haemolyticus* uncovers

RT extreme plasticity of its genome and dynamism in the evolution of  
 RT human-colonizing staphylococcal species."  
 RL Submitted (PRC-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP006716; BA00520.1; - ; Genomic DNA.  
 SQ SEQUENCE 1361 AA; 148706 MW; D3D4A61912P4998B CRC64;  
 Query Match 6.7%; Score 119; DB 2; Length 1361;  
 Best Local Similarity 20.8%; Pred. No. 13;  
 Matches 90; Conservative 58; Mismatches 143; Indels 142; Gaps 22;  
 QY 2 TLRDMSMDTG-VSTFNVLTETGLDNCTNLYANGLNMGIVINITPTDDEGNFVDIDVT 60  
 Db 721 TLKISNLKOTILGQVSTKNGKTYTTVTDKVKANA-TIGGKVNIT-----KGAT 768  
 Qy 61 LNDNDIKIVDYGSDIDG--SGDW----- 92  
 Db 769 LNNK---DYYLISDYNGNTSRGVNPASETSKVTTTPNKVNATYDINGNANIYDTPWGT 824  
 Qy 93 - ----TINS---QSYSLKSEN-----QIIOIKRIVSCSVTSRLRK 128  
 825 AKOVKATVPSGKQKLSIDPNLKVGENETYLHGVINNIWGNIKSTDVKQTPKVNAPAVNTK 884  
 Qy 129 SF--SAKUTTSGKVVISITONSINSRRVINAL--DATMFTDDLERTRKTRPENQSTS 184  
 Db 885 ATKKSTASQTQTVSKVQAQWNANGARATYDKTKGNATKVNRFNVSGETIDGNTVL 944  
 Qy 185 HKSSSTNSLVYHTWTPRSLKQN-----WRWEDTYNG----WTWAQ----- 221  
 Db 945 LQNTTTNTPL-GWFENVKDLKVKNSSSEOKTSGTYKINNNKNGLYSIAMGCTNQQLDNLKD 1003  
 Qy 222 -----SCYYKKGADGSESTRWLAQSI-----PPPGMWD 251  
 Db 1004 ANKSFKVKSVWIGNTTYLGTVNG--KTCGWIAKNDLTSSNTISSDNGEKYQYEFLINTN 1061  
 Qy 252 GLWLNDNDIALSGMAHKSINYDTGINOLSFTRIIGKGFSAWYNISQLDRGHAVIID--- 307  
 Db 1062 GYYVDDPSSAKATSUKAFNEQ--IFQVTRKVV-NGKTMWY--GKLNSQKVWIKDTLK 1116  
 Qy 308 ---OYGNKVRIL 316  
 Db 1117 KELVKVSKSYRTL 1129

Search completed: January 30, 2006, 09:48:54  
 Job time : 29.6268 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model  
Run on: January 30, 2006, 09:40:03 ; Search time 128.372 Seconds  
(without alignments)  
Scoring table: BLOSUM62

Title: US-09-889-874A-23  
Perfect score: 8879  
Sequence: 1 VVVKPLKLFRRTWSDNNEP.....PRKILGRTTEKTVKPKTFRP 1673  
5726.190 Million cell updates/sec

Searched: Gapop 10.0 , Gapext 0.5  
2443163 seqs, 439378/81 residues

Total number of hits satisfying chosen parameters: 2443163  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseg\_21:  
1: genesegp19808:  
2: genesegp19908:  
3: genesegp20008:  
4: genesegp20108:  
5: genesegp20028:  
6: genesegp20038:  
7: genesegp2003bs:  
8: genesegp2005bs:  
9: genesegp20058:  
\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description	ALIGNMENTS
1	8879	100.0	1673	3 AAY5707 Aay95707 Cosmid ch Aay33727 Photorhab Ad26337 Chromobac Adm67433 Photorhab Adm67283 Photorhab Adg73113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	RESULT 1 ID AAY95707 standard; protein; 1673 AA. XX AC XX AAV95707; DT 25-OCT-2000 (first entry) DE Commid CHRIM5 encoded protein PI4-2F. XX KW Cosmid CHRIM5; nematocide; nematode; biological control agent; KW transgenic plant; helminthiasis; PI4-2F. XX OS Xenorhabdus bovienii. XX PN W02000-2855-A1. XX PD 27-JUL-2000. XX PF 24-JAN-2000; 2000WO-GB000219. XX PR 22-JAN-1999; 99GB-00001499. XX (HORT-) HORTICULTURE RES INT. PT Morgan JAW, Jarrett P, Ellis D, Ousley MA; XX DR WPI: 2000-49157/44. N-PSDB; AAS0029. XX PT Novel composition used to control parasitic nematodes, especially in a plants such as maize, cotton, soya, and rice, comprises a bacterium which is a symbiont of an entomopathogenic nematode. XX PS Example 6; Page 42-43; 74pp; English.
2	1638	18.0	1584	2 AAY33727 Aay33727 Photorhab Ad26337 Chromobac Adm67433 Photorhab Adm67283 Photorhab Adg73113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	Abn69806 Photorhab Abu40318 Protein e Abn6973 Photorhab Abu16634 Protein e Abc79678 Pseudomon Abu22662 Protein e Abd3317 Acinetobac Abu16693 Protein e Abu4051 Protein e Adi05044 Bacterial Abu21384 Protein e Aau5772 Propionib Abm49291 Propionib Abm65127 Propionib Abu50462 Protein e Adi07275 Bacterial Abc75059 Pseudomon Abu4090 Protein e Adi16955 Murine NO Adi16953 Chicken N Abu47385 Protein e
3	1601	18.0	1410	8 ADQ26337 Ad26337 Chromobac Adm67433 Photorhab Adm67283 Photorhab Adg73113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	Abn69806 Photorhab Abu40318 Protein e Abn6973 Photorhab Abu16634 Protein e Abc79678 Pseudomon Abu22662 Protein e Abd3317 Acinetobac Abu16693 Protein e Abu4051 Protein e Adi05044 Bacterial Abu21384 Protein e Aau5772 Propionib Abm49291 Propionib Abm65127 Propionib Abu50462 Protein e Adi07275 Bacterial Abc75059 Pseudomon Abu4090 Protein e Adi16955 Murine NO Adi16953 Chicken N Abu47385 Protein e
4	1583	17.8	1787	6 ABM67433 Adm67433 Photorhab Adm67283 Photorhab Adg73113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	Abn69806 Photorhab Abu40318 Protein e Abn6973 Photorhab Abu16634 Protein e Abc79678 Pseudomon Abu22662 Protein e Abd3317 Acinetobac Abu16693 Protein e Abu4051 Protein e Adi05044 Bacterial Abu21384 Protein e Aau5772 Propionib Abm49291 Propionib Abm65127 Propionib Abu50462 Protein e Adi07275 Bacterial Abc75059 Pseudomon Abu4090 Protein e Adi16955 Murine NO Adi16953 Chicken N Abu47385 Protein e
5	1569	17.7	1590	6 ABM67433 Adm67433 Photorhab Adm67283 Photorhab Adg73113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	Abn69806 Photorhab Abu40318 Protein e Abn6973 Photorhab Abu16634 Protein e Abc79678 Pseudomon Abu22662 Protein e Abd3317 Acinetobac Abu16693 Protein e Abu4051 Protein e Adi05044 Bacterial Abu21384 Protein e Aau5772 Propionib Abm49291 Propionib Abm65127 Propionib Abu50462 Protein e Adi07275 Bacterial Abc75059 Pseudomon Abu4090 Protein e Adi16955 Murine NO Adi16953 Chicken N Abu47385 Protein e
6	932.5	10.5	982	7 ADG3113 Adg3113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	Abn69806 Photorhab Abu40318 Protein e Abn6973 Photorhab Abu16634 Protein e Abc79678 Pseudomon Abu22662 Protein e Abd3317 Acinetobac Abu16693 Protein e Abu4051 Protein e Adi05044 Bacterial Abu21384 Protein e Aau5772 Propionib Abm49291 Propionib Abm65127 Propionib Abu50462 Protein e Adi07275 Bacterial Abc75059 Pseudomon Abu4090 Protein e Adi16955 Murine NO Adi16953 Chicken N Abu47385 Protein e
7	932.5	10.5	982	7 ADG3113 Adg3113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	Abn69806 Photorhab Abu40318 Protein e Abn6973 Photorhab Abu16634 Protein e Abc79678 Pseudomon Abu22662 Protein e Abd3317 Acinetobac Abu16693 Protein e Abu4051 Protein e Adi05044 Bacterial Abu21384 Protein e Aau5772 Propionib Abm49291 Propionib Abm65127 Propionib Abu50462 Protein e Adi07275 Bacterial Abc75059 Pseudomon Abu4090 Protein e Adi16955 Murine NO Adi16953 Chicken N Abu47385 Protein e
8	517.5	5.8	1317	8 ADG26337 Ad26337 Chromobac Adm67433 Photorhab Adm67283 Photorhab Adg73113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	Abn69806 Photorhab Abu40318 Protein e Abn6973 Photorhab Abu16634 Protein e Abc79678 Pseudomon Abu22662 Protein e Abd3317 Acinetobac Abu16693 Protein e Abu4051 Protein e Adi05044 Bacterial Abu21384 Protein e Aau5772 Propionib Abm49291 Propionib Abm65127 Propionib Abu50462 Protein e Adi07275 Bacterial Abc75059 Pseudomon Abu4090 Protein e Adi16955 Murine NO Adi16953 Chicken N Abu47385 Protein e
9	432	4.9	2334	5 ADG3113 Adg3113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	Abn69806 Photorhab Abu40318 Protein e Abn6973 Photorhab Abu16634 Protein e Abc79678 Pseudomon Abu22662 Protein e Abd3317 Acinetobac Abu16693 Protein e Abu4051 Protein e Adi05044 Bacterial Abu21384 Protein e Aau5772 Propionib Abm49291 Propionib Abm65127 Propionib Abu50462 Protein e Adi07275 Bacterial Abc75059 Pseudomon Abu4090 Protein e Adi16955 Murine NO Adi16953 Chicken N Abu47385 Protein e
10	386	4.3	2234	6 ADU18641 Abu18641 Protein e;	
11	365	4.1	1400	7 ADC01365 Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	
12	361	4.1	1510	7 ADG3113 Adg3113 Pseudomon Adl12168 Pseudomon Adg22210 Bacterial Alg31849 Human kin Abu18641 Protein e; Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	
13	353.5	4.0	843	6 ADU18641 Abu18641 Protein e;	
14	350	3.9	1515	6 ADU18641 Abu18641 Protein e;	
15	342	3.9	1397	6 ADU18641 Abu18641 Protein e;	
16	341.5	3.8	1404	6 ADU18641 Abu18641 Protein e;	
17	341	3.8	1394	7 ADC01365 Adc01365 Enterohae Adg30698 Xanthomon Abu48290 Protein e Abu19676 Protein e Abu15135 Protein e Adc00960 Enterohae Adc01426 Enterohae Abu15136 Protein e Abu15983 Protein e Abu14693 Protein e Abu34791 E. coli c Aag9897 E. coli g Abu14807 Protein e Abu14191 Protein e	
18	340.5	3.8	1411	6 ADU18641 Abu18641 Protein e;	
19	339	3.8	1426	3 ADU18641 Abu18641 Protein e;	
20	339	3.8	1426	6 ADU14693 Abu14693 Protein e;	
21	331.5	3.7	1377	4 ADU14693 Abu14693 Protein e;	
22	331.5	3.7	1377	4 ADU14693 Abu14693 Protein e;	
23	331.5	3.7	1377	4 ADU14693 Abu14693 Protein e;	
24	331.5	3.7	1377	4 ADU14693 Abu14693 Protein e;	
25	328	3.7	1572	6 ADU14693 Abu14693 Protein e;	

The present sequence is that of protein PI4-2F encoded by an open reading frame identified in cosmid CHRIM5 (see AAS0029). CHRIM5 was obtained by ligating Xenorhabdus bovienii strain 173 (NCIMB 40986) Sau3A-digested DNA fragments into the BamHI site of the Stratagene cosmid vector SuperCos1, packaging into Escherichia coli XL Blue 1, and screening for nematicidal activity against Caenorhabditis elegans. Analyses of the DNA indicated a number of open reading frames for which the corresponding protein sequences were determined (see AAS0685-V95735). Nematodes can be controlled through the use of bacteria associated symbiotically with an entomopathogenic nematode. Such bacteria include Xenorhabdus and Abi15136 Protein e  
Abi15983 Protein e  
Abi14693 Protein e  
Abu34791 E. coli c  
Aag9897 E. coli g  
Abu14807 Protein e  
Abu14191 Protein e  
Abu48290 Protein e  
Abu19676 Protein e  
Abu15135 Protein e  
Adc00960 Enterohae  
Adc01426 Enterohae  
Abu15136 Protein e  
Abu15983 Protein e  
Abu14693 Protein e  
Abu34791 E. coli c  
Aag9897 E. coli g  
Abu14807 Protein e  
Abu14191 Protein e  
Abi15317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385 Protein e  
Abu40318 Protein e  
Abn6973 Photorhab  
Abu16634 Protein e  
Abc79678 Pseudomon  
Abu22662 Protein e  
Abd3317 Acinetobac  
Abu16693 Protein e  
Abu4051 Protein e  
Adi05044 Bacterial  
Abu21384 Protein e  
Aau5772 Propionib  
Abm49291 Propionib  
Abm65127 Propionib  
Abu50462 Protein e  
Adi07275 Bacterial  
Abc75059 Pseudomon  
Abu4090 Protein e  
Adi16955 Murine NO  
Adi16953 Chicken N  
Abu47385

CC an engineered bacterium, or a nematocidal protein obtained from such  
 CC bacteria, particularly p13-1f (see AAY95706) or p14-2f can be used to  
 CC control helminths in a human or domesticated animal or for the  
 CC control of plant pathogen nematodes. Also claimed are vectors for  
 CC expressing nematocidal proteins in host cells, and transgenic plants  
 XX

## Sequence 1673 AA:

Query Match	100.0%	Score 8879;	DB 3;	Length 1673;	
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;
Matches 1673;	Conservative				
QY	1	VYIKELKUERRITMSDNNEPPTQANNTSAVSGCDPRTGLMNIQITGHINGNTGPT	60		
Db	1	VYIKELKLUERRITMSDNNEPPTQANNTSAVSGCDPRTGLMNIQITGHINGNTGPT	60		
QY	61	LPLTISYSPLNKTDIGFGIFGNFGLSVYDRKNSLISLSTGENYKVIEDKTVKLQOKKD	120		
Db	61	LPLTISYSPLNKTDIGFGIFGNFGLSVYDRKNSLISLSTGENYKVIEDKTVKLQOKKD	120		
QY	121	NLRFEKDLCENCYRIRHKGDEVLTLGEGNNNAFLDKVPLKLNPGAHAYTIDNPEATQP	180		
Db	121	NLRFEKDLCENCYRIRHKGDEVLTLGEGNNNAFLDKVPLKLNPGAHAYTIDNPEATQP	180		
QY	181	RLRKTYDDGGHDPLNTEYQGLKTIILTPCQEGRTEFLNFOLNTHFESGN	240		
Db	181	RLRKTYDDGGHDPLNTEYQGLKTIILTPCQEGRTEFLNFOLNTHFESGN	240		
QY	241	ENPLTWFSFGTIPIGKNGILQQWITSMTAPGKLTETVNISSANNQHHPPOSANLPVLPVT	300		
Db	241	ENPLTWFSFGTIPIGKNGILQQWITSMTAPGKLTETVNISSANNQHHPPOSANLPVLPVT	300		
QY	301	LMKOVPGAGAOPAIQAEYTSIHHYVGGSNGIWNKKNLDNLIVGLMEINYGTSERRYOK	360		
Db	301	LMKOVPGAGAOPAIQAEYTSIHHYVGGSNGIWNKKNLDNLIVGLMEINYGTSERRYOK	360		
QY	361	EHGHDQIVRLTERTYNNHLLTSECKQNGYIOTTEAYVAILGHNFDSQSFOLPKTKE	420		
Db	361	EHGHDQIVRLTERTYNNHLLTSECKQNGYIOTTEAYVAILGHNFDSQSFOLPKTKE	420		
QY	421	TWRSADNSYSESBITETFDSSGNPLTKVTKDCKTOKIISPHTHEWYPAGEVDNCPP	480		
Db	421	TWRSADNSYSESBITETFDSSGNPLTKVTKDCKTOKIISPHTHEWYPAGEVDNCPP	480		
QY	481	YGFTTRPVKKLIQTDPSEKFKEPKIQYRSLIGSOSHWTKEERHYSATOLNSTLF	540		
Db	481	YGFTTRPVKKLIQTDPSEKFKEPKIQYRSLIGSOSHWTKEERHYSATOLNSTLF	540		
QY	541	QYNTOKSELGRLLKOPCTKEENGKYSVWVKPFTYKQDQLQSHSITTHDNFTHSQ	600		
Db	541	QYNTOKSELGRLLKOPCTKEENGKYSVWVKPFTYKQDQLQSHSITTHDNFTHSQ	600		
QY	601	VRSRTYGRLFSDPTDKDIDVOMSYDKLGRJLRTJNSGTYANTITYDELLNQDDNRP	660		
Db	601	VRSRTYGRLFSDPTDKDIDVOMSYDKLGRJLRTJNSGTYANTITYDELLNQDDNRP	660		
QY	661	PFVITTDVNQNLNRNEFDGIGRHYSOCLKXSDGDGKFYHTHQOYDECGRHHTSYSDY	720		
Db	661	PFVITTDVNQNLNRNEFDGIGRHYSOCLKXSDGDGKFYHTHQOYDECGRHHTSYSDY	720		
QY	721	LTNGROQTDPKVHLSMSKSTDNNGQIANTHWSGYSEKITYDPTLTAKQLOSNNNV	780		
Db	721	LTNGROQTDPKVHLSMSKSTDNNGQIANTHWSGYSEKITYDPTLTAKQLOSNNNV	780		
QY	781	QTGKEETIYTQSQPOQITLDEAGHLOSCILTRGDWDRKEDTAIGCTIQDNTN	840		
Db	781	QTGKEETIYTQSQPOQITLDEAGHLOSCILTRGDWDRKEDTAIGCTIQDNTN	840		
QY	841	RVQITLPDGTVNRYKXAPFSTDTLDIRNGISLGQQFDGLSRITOSQDGRWVAYT	900		
Db	841	RVQITLPDGTVNRYKXAPFSTDTLDIRNGISLGQQFDGLSRITOSQDGRWVAYT	900		
QY	901	YASAGNDQCPSTVTPDQGQFIHVQYQPELDDAVLQVASNETITQFSYNPVTCALLKAVAEG	960		
RESULT 2					
	AAY3727				
ID	AAY3727	standard; protein; 1584 AA.			
XX					
AC	AAY3727;				
XX					
DT	0-NOV-1999 (first entry)				
XX					
DE	Photorhabdus luminescens 176 kd insecticidal toxin.				
XX					
KW	Symbiotic bacterium; nematode; insect; larva; toxin; insecticide.				
XX					
OS	Photorhabdus luminescens.				
XX					
PN	W09942589-A2.				
XX					
PD	26-AUG-1999.				
XX					
PP	18-FEB-1999; 99WO-EP001015.				
XX					
PR	20-FEB-1998; 98US-00027080.				

PR XX 20-JAN-1999; 99US-0116439P.  
 PA (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERRINDINGEN VERW GES MBH.  
 XX  
 PI Kramer VC, Morgan MK, Anderson AR, Hart HP, Warren GW, Dunn MM;  
 PI Chen JS;  
 DR WPT; 1999-5227479/44.  
 DR N-PSSDB; AAZ06026.  
 XX  
 PT New nucleic acid from *photorhabdus luminescens* encoding insecticidal  
 toxins, used for making resistant transgenic plants.  
 XX  
 PS  
 XX  
 Claim 25, Page 93-98, 148pp; English.  
 This sequence represents an approximately 176 kD insecticidal toxin from  
*Photorhabdus luminescens*. It is one of five insecticidal toxins (AAU33723;  
 CC -Y33727) encoded by open reading frames (ORFs) arranged in an operon-like  
 structure in a 9.7kb fragment of *P. luminescens* DNA (AAU60826). This  
 sequence is encoded by ORF5. *P. luminescens* is a member of the  
*Enterobacteriaceae* family and is a symbiotic bacterium of nematodes of  
 the genus *Heterorhabditis*. The nematodes colonise insect larvae, kill  
 them, and their offspring feed on the dead larvae. However, the  
 insecticidal agents are produced by *P. luminescens* rather than the  
 nematodes. The toxins have activity against Lepidopteran insects such as  
 Cabbage Looper (*Trichoplusia ni*), European Corn Borer (*Ostrinia nubialis*)  
 and Fall Armyworm (*Spodoptera frugiperda*) and also against Coleopteran  
 insects (e.g., Colorado Potato Beetle, *Lepinotarsa decimlineata*). In  
 addition the toxins are active against strains resistant to known  
 insecticides. The DNA sequence can be used to generate transgenic plants  
 of various species that are resistant to economically important insect  
 pests and also for recombinant production of the toxins for use as  
 insecticides  
 XX  
 SQ Sequence 1584 AA;  
 Query Match 18.4%; Score 1638; DB 2; Length 1584;  
 Best Local Similarity 31.0%; Pred. No. 4.6e-105;  
 Matches 493; Conservative 233; Mismatches 682; Indels 182; Gaps 54  
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 19 EFFTQANNFTSAVSGQDVPRGQLYNTQITLGHIVGN--GNQOPTPLTLYSPNLTNSPLNKTDIG 76  
 5 DIVSNAPNFQSYINTVDPRTQOYSANINITIRPNVNGNSQT--LSLSFSPLTTLNG 62  
 77 FGICGRNFGSLSVYDRKNSLSLSTGENYK--VIETDKTKWLUQOKKUDNLREFKDLENKY 133  
 63 FGIGKRFLSLTDLKILTFSRANGSDFPKCKLFLPPNNNDSFKDKKKUDLVRK-LDSNTF 121  
 134 RIIHKGSDIEVLTGTMNAFDLKVPKKLUNPAGHAYIDMNPFATQPRLNRITYDDLDHD 193  
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 122 YVNKNRNGIBILKRIGSS--DIAKTVALEFPDGEADLIVNSR----- 162  
 194 IPILNLEYQGLIKITLT--EPPGQKSCYRTERLFL-NRQLNSIHNFSLGENENFLTWSIGTY 251  
 163 FALSEKIVKVTGKTVKLNSGNN--CTSEYPPDNNSIKAFDYRNDLVLITVFD 219  
 252 PIGK-----NGTIGQWITSMTAGGLEKETVNYNNQHHPPOSANLPYLY 298  
 220 ASGPIDSARFKMTIQTLKGKF--PVISTFRTPGYELVSYKEN--GH--KVTDBEVY 273  
 Qy :|||:|||:|||:|||:|||:|||:|||:  
 Db 299 VTLMKKQVPGAGQPAQIAQEYTS-TNYVGGCGNGL-WNNKLDNLQIGLMTEYNGSRSRR 356  
 274 AAALTIQPGENGQPAVSKSYVSSVNFGLYSSGRTSFDSODDNLYLUTGKYSTIE-- 330  
 Qy :|||:|||:|||:|||:|||:|||:  
 Db 357 YKQKGHDQVRIERETTYYNHLTBCKQONGYIQTETAYAIGHNFDSOPQSQFOLP- 415  
 331 -RVLDGQSVVSVIERNVKHMLTMTEAKTOQNKRITTEITNEDLKSFSQSPRENTQOQS 389  
 416 KTKTETWRSADNSYSEITEETTDFDSGNPLTKVTKRKTKTIPSPTHMVEYPPAGEVDN 475  
 Qy :|||:|||:|||:|||:|||:  
 Db 390 RVLTRYDQINTSBETWIKSDMGTNTL-ITETSGIQ----EIVVYPPNGBGS 442

Db	1464 SQAVSAGVIGSVPLPF----GEVASRSSL 1488
RESULT 3	
ID ADQ26337	ADQ26337 standard; protein; 1410 AA.
XX	
AC ADQ26337;	
XX	
DT 23-SEP-2004	(first entry)
XX	
DE Chromobacterium violaceum CV00361 involved in hydrocyanic acid synthesis.	
XX	
KW Hydrocyanic acid; biocide; pesticide; fungicide; antibacterial; nematocide; CV00361.	
XX	
OS Chromobacterium violaceum.	
XX	
PN WO2004056960-A2.	
XX	
PD 08-JUL-2004.	
XX	
PR 16-DEC-2003; 2003WO-BR000207.	
XX	
PA (CNPQ-) CNPQ CONSELHO NACIONAL DESENVOLVIMENTO.	
XX	
PI De Vasconcelos ATR, Simpson AUG, Abreu HNS, De Almeida DF;	
PI Almeida FC, De Almeida R, Antonio RV, Araripe JR, De Arruda MFP;	
PI Bogo RM, Bonatto SL, Brígido MM, De Brito CPA, Bruchi M, Buriti HA;	
PI Camargo A, Carraro D, Carvalho CMB, Cascudo JCOM, Cavada BS;	
PI Chueire LMD, Da Cunha MH, Fantinatti F, Farias IP, Felipe MSS;	
PI Ferrari LP, Ferro JA, Franco GR, De Freitas NSA, Furlan LR;	
PI Gattapaglia D, Gazzinelli RT, Gomes JAA, Goncalves PR, Grangeiro TB;	
PI Gribaldo SC, Guimaraes CT, Hanna ES, Jardim SN, Laurino JP, Lima LFA;	
PI De Lyra MDCCP, Madeira HMF, Maiamhao AQ, Manfio GP, Martins WS;	
PI De Medeiros SRB, Meissner RDV, Moreira MA, Do Nascimento FF;	
PI Nicollas MF, De Oliveira JG, Oliveira SC, Paixao RFC, Parra J;	
PI Pasa TBC, Pedrosa FDO, Pena SDP, Pereira JO, Peleira M, Pinto LSRC;	
PI Pinto LDS, Porto JIR, Potrich DP, Ramalho CE, Reis AW;	
PI Rondinelli E, Sampayo AH, Dos Santos FR, Schneider MEC, Silva DW;	
PI Silveira R, Soares CHDA, De Souza EM, Souza RC;	
PI Steindel M, Teixeira SMR, Trevilato PB, Urmenyi TP, Wassen R;	
PI Azevedo V, Bartolleti LA, Batista JDS, Filho AS, Zaha A;	
PI Andrade EDRM, Gonçaga LA, Dos Santos EBP, Soares RDBA, Batatas LAM;	
PI Cardoso DDVDP, Parente JA, Rigo LU, Steffens MBR;	
XX	
DR WPI, 2004-500292/47.	
XX	
PT New gene-coding polynucleotides of the chromosome of Chromobacterium violaceum, useful for therapeutic, diagnostic or pharmacological applications, in control processes for environmental parameters or in enzyme synthesis.	
XX	
PS Claim 2; SEQ ID NO 10; 31PP; English.	
XX	
The present sequence is that of the protein encoded by the CV00361 gene of Chromobacterium violaceum strain ATCC 12472 (NCIB 9111, NCTC 9757, JCM 1249, DSM 30191, IAM 12470, IMG 1267). The invention relates to 29 polynucleotides that have been identified by sequencing the genome of this strain, to the polypeptides ADQ26328-ADQ2635 encoded by these polynucleotides, and to the uses of the polynucleotides and polypeptides for therapeutic, diagnostic, medicinal, pharmacological and pharmacognostic applications, in control processes for environmental parameters, and in enzyme synthetic processes. The CV00361 gene is involved in the synthesis of hydrocyanic acid. The gene and the encoded polypeptide can be used in the generation of agricultural biocides useful against bacteria, fungi and nematodes.	
XX	
Sequence 1410 AA:	





QY 180 PRILNRTYDLDGHDIDPLNTLEYQGLIKITLT-FPGOKEGVTELRL-NRQLNSIHFS 237  
 Db 117 ---EVFDLYINSRPALSEKIVRTGKTYLKNVSGNN--CTSVEPDNNNISAKTAFD 169  
 Qy 238 LAMENPLTWSFGCPTIGK---AGLGMWISTMSTAPGELKETTINYSNNG 284  
 Db 170 YRNDDYLITWTVPYDASGPIDSARFKMTYOTLKGF-PVTSARFTPTGVYELVSYKEN-G 226  
 Qy 285 HHRPOSANLPVLPVTLMKQVPGAGQPAIQAEVSYTS-HNYVGGGSNGI-WNKLDLYG 342  
 Db 227 HKYTDTESIPYAAALIQ-PANGQFRIKSKEYSSTANPQYSSORTSFESSQDNYL 283  
 Qy 343 LMTEYNKGSTERSRYKDGEGHDOIVRERTTYNNYHLLSECKQONGYIOTTAYAHLG 402  
 Db 284 VICKYTTSIE---RVLINGONVISUTERVFELKFLHMWKTEAKHQDNKRITITYNEDPS 339  
 Qy 403 KNSFSEOPENLOQSHVLTTRYTDLT---TKETMSADNSRSETETTEDBESGVLTKIKDKTQ 456  
 Db 340 KSFSEOPENLOQSHVLTTRYTDLT---NTSREBESVNIAKDDWGNTL-TETSGIOK 393  
 Qy 457 IISPSTHWYYPAGEVNCPPBYGFRPKVKLIQTPYDSFKDPDEKFQYRSLIGS 516  
 Db 394 ---EVYYPRNGEBCNGCPADLGFRKLKVSTQKSPDAQSVARVNTSYTYQKPT 447  
 Qy 517 QSHVTLKIEERHYSATOLNST---ROY---NTDKSELGRJIKQETCTKGNGKTVSW 570  
 Db 448 FTGAYVK---EYVSKASETIDSKVURTFENYVNSPNTKSH-GSLAKITSMNO---QTV 499  
 Qy 571 HRTFTYKDDTTLQOSHSTTHMPTIHSQHRSVRYTRGBLSPDMDKTOVITOMSYDKGRL 630  
 Db 500 TTPKVEYDSEMTNSTVTFGDFGTHMSEKVNVIYTHRLRKVDNVTIDOSYDLSGR 559  
 Qy 631 LRTTLNLSGTPYANTLTQYELNNLQDDNRPPFVTTDVNGQNLNEFDRGAGRHVSOCLK 690  
 Db 560 ISQILDCTKETKRSVYIQYQCGDENDFWP-VNIEDSQGRRKTHDMGRICSEEQ 618  
 Qy 691 DSDG-----DGKFYHTHTQYDEOCRHHYSTSYDLTN---GROQTDPDKVILM 737  
 Db 619 DDDGVWGTSGIKOGTYKULARQYDVLGOLVKERISNDLWDISANPLTRLTIP-LVTK 676  
 Qy 738 SISYDNMGQIANTHWSYKUSEKTVDPTLTAK---QLOSNMNQVOTKEVNTYPT 791  
 Db 677 TXQYDGWGNRYSTBSPDRIBELBHDPTRTTQHGKGIGMLNITQNNP----- 725  
 Qy 792 SQOPIQITPDEAGLQSCHTLTDGMDVRKETDALGOCTYQYDINVRVQITLGD 851  
 Db 726 -EOPASIKVYVPGAIYSTRTRYDGEERTVYDADSYATQEYDIEFDRTKKTLDT 784  
 Qy 852 TMRKVKARPFSTMLTIDRVNGISLGGOPTDGLSLRQTSODS3CRVWAVTYSGNDOCST 911  
 Db 785 IESAYASFSHEBLISALVNQGOLGLSVLWYDGLGRVTRDTCERKTEVLYGSGDK-PIQ 843  
 Qy 912 VITP--DQFIRHQYQPELDADDLQVSNNETIQFSTPNTVPTALKAVAGSOLPTI-YV 958  
 Db 844 SVTPAHNQKNTDLY--ALGSVNSKFTETTSOONPSYQTKGALLSA-TEGVSYNSYF 900  
 Qy 969 PSGRKJEMENIN-DMKKGS---YVLTRELENGSTDLTQIKISRDTHGRVTKQKDSIK 1024  
 Db 901 PSCVQHRSRSPRINKPISSGDQYTMSSGLQHKSFDHRYHYSYDEGRVYKTEQSOY 950  
 Qy 1025 TUNYDDLNRRHGSQVDTLATEGMLLTVEPDGLNREIGKCDSSGHTDQOSWLKTO 1084  
 Qy 961 ATPEFYDVGRLITTTKDTTSQSLATKIEYDVFDRKSLISDFSIQV-ITLSYTKAN 1019  
 Qy 1085 QLARNRIVLNGIYQRTSOSYNSRNLMQYQKDGAECPDQKXHISTVNTFEDYKNT 1144  
 Db 1020 QISQRITSIDGVNMKNERQYQISNORLQYQCEQBDQVHTRVLSOOVHYDOWENIK 1079  
 Qy 1145 ACHTFADGDEDHATPKKANPDPQCQLTEVHHTHPMDMNLKUYDKRAGRVINIT-DHNG 1203  
 Db 1080 RLQNTYRSGKET-VDHSQ-AUPTOLRI---TSDQKQIISYDANG--NLTRBKG 1130

Qy 1204 NTENFTVTDLGL---QNGQESV--YGYDPNLRLVSQL--KTDLDCELYTRETMVNEV 1255  
 Db 1131 QT-LIXDQNNRNLVQVGDGSKGNLVCQYDANKLTAQVLANGTVRQ-YASGNVANQ 1187  
 Qy 1256 RNCGEMTIL-----RTGETTIAQOBASKVUQITGTDQOSVLTEDKQNLQSOA 1303  
 Db 1188 LGDETTIWLSQDKQRUGHQSTKNGESVYQY-----GTDHNSTVIASONEELMAIS 1239  
 Qy 1304 YSAYGKHKSTANDASTIGNGERADPVGVTIHLNGYRVSQPTLMRHTPSLSPFGAGG 1363  
 Db 1240 YPYGFRSLI--SSLPGLNGAQDVPVGMWFLGNGYRVENFVPLMPHSPDWSWPFCRG 1296  
 Qy 1364 IMPYCYGDPINRSDDSGHSIWAQWTGIGMGIAGLUMIAT--GENMIA---AACGIA 1417  
 Db 1297 VNPYTYCGDGPINRIDINGHLSAGGLGIVLGAIGIIVGVUSLGAAGAISAGLIAAGGAL 1356  
 Qy 1418 AJASTSTTALFGALSUTSDTISVSGALDASPKRASLNGWVSMGAGLAEST-- 1475  
 Db 1357 GAIASTSRAFAVATVIGLAADSIGASALSEKDPKVAGLAWNISTIGLVLSFGISAITF 1416  
 Qy 1476 -----KGGKLATH-LGAF-AEDGENALKTSESSRIKWMTRS 1513  
 Db 1417 TESLIKSRASSQSVASTSVISVSPFEGEA----SSRR-RWDIALS 1459

RESULT 6  
 ID ADG73113 standard; protein; 982 AA.  
 XX ADG73113;  
 XX DT 11-MAR-2004 (first entry)  
 DE Pseudomonas syringae pv. tomato DC3000 Avr/Hop protein #53.  
 XX KW Avr; Hop; transgenic plant; disease resistance; cancer; bacteria; metabolic pathway; eukaryotic cell death; programmed cell death; cytostatic.  
 XX OS pseudomonas syringae; pv. tomato str. DC3000.  
 XX PN US2003204868-A1.  
 XX PD 30-OCT-2003.  
 XX PP 12-FEB-2003; 2003US-00365742.  
 PR 12-FEB-2002; 2002US-0356408P.  
 XX 10-MAY-2002; 2002US-0380185P.  
 PA (COLL/) COLIMER A.  
 PA (ALFA/) ALFANO J R.  
 PA (CART/) CARTINHOUR S W.  
 PA (SCHN/) SCHNEIDER D J.  
 PA (TANG/) TANG X.  
 XX PT Colomer A, Alfano JR, Cartinhour SW, Schneider DJ, Tang X;  
 XX DR WPI: 2003-875735/81.  
 DR N-PSDB; ADG73112.  
 XX New nucleic acid, useful in imparting disease resistance to a plant or in preparing a composition for treating cancer.  
 XX encoding them. Also disclosed are expression vectors, host cells, and transgenic plants comprising polynucleotide sequences of the invention.  
 CC The polynucleotide and polypeptide sequences are useful in imparting disease resistance to a plant or in preparing a composition for treating cancer. The sequences may also be used to make a plant hypersusceptible

CC	to colonisation by nonpathogenic bacteria, modify a metabolic pathway in	DB	852 PVGLVGRGAGAVVARAVAAKVAANAGTIGRSVSRIG-LAAGARRTSSAASSR 909
CC	a cell, cause eukaryotic cell death, and inhibit programmed cell death.	CC	
CC	The present sequence represents a Pseudomonas syringae pv. tomato DC3000	CC	
XX	Avr/Hop protein of the invention.	XX	
SQ	Sequence 982 AA;	RESULT	7
Query Match	10.5%; Score 932.5; DB 7; Length 982;	ID	ADL12168 standard; protein; 982 AA.
Best Local Similarity	29.7%; Pred. No. 8.3e-56;	AC	ADL12168;
Matches	285; Conservative 142; Mismatches 357; Indels 175; Gaps 33;	XK	
QY	642 ANTYDYEELNNLQDDNRPFPVTTDNGNQRNEFGAGRAVSQLKDSDGDKFVIT 701	DT	06-MAY-2004 (first entry)
Db	31 STLGYRD----DWNR--CCITTDDN-VQTYEYSPIGSDVHK-----GPIKTT 74	XK	Pseudomonas syringae anti-cancer protein #53.
QY	702 HTQOYDEQGRHHTSTYSYLNGRQTPDPKVHLMSKSYDNMQIANTHWSYGVSEKIT 761	OS	KW cytostatic; gene therapy; Avr; Hop; cancer.
Db	75 WKQSGDPBGR-----ISGSET-----WNLN-----RGPDRIR 103	XK	Pseudomonas syringae; pv tomato DC3000.
QY	762 VDPITLTATKQLOQNNSNNVQTGKEVTTIPSPQQPIQITLPDEAHLQSCHTLRDGMWRV 821	PN	PN W02003168930-A2.
Db	104 ---TLAGTKGRSRTHMSRSRLTT--TEQL-----SRQFLYDGLRC 144	XX	
QY	822 RKETDAIGCCTIQYDNYKRVIOITLPGTTVNPNEYKAPEPSTDTLTDIYV--NGIS-- 875	PR	PR 10-MAY-2002; 2002US-0380165P.
Db	145 TEORDALQOSTLFSYDNNSRMVSVSSTLADGSVINSYAPLOSSELATMLLEVHONGTRIV 204	PD	PD 21-AUG-2003.
QY	876 LGQQTFDGLSLRRTOSQDPGGRWAVTYSAGNDQCPSTVTPDGQPHYQXQPBDLAVQ 935	XK	XK 12-FEB-2003; 2003WO-US004450.
Db	205 AGTOKFDGLERVTQTKTGVRVEQFYDAGEMQ-PRSRITAGLDNINFYTRALDQISS 263	PA	PA (CORR ) CORNELL RES FOUND INC.
QY	936 ASNETTQFSYKNPVIGALKAVAEQSLTPYYPSGRGLKMNINDKMSYLM-----T 989	PA	(USDA ) US SRC OF AGRIC.
Db	264 TAPDETAKEDYKTSARLLETAINFOGTRTRVYDHNLQGETWNL--LGQAMETRHQS 321	XK	(UNIV ) UNIV NEBRASKA.
QY	990 LRGLENGYDIL---TGTIOKISRDTHRGTQVTDQKDSSTKTTINYDLDLRHQSQVTDIA 1044	PA	(UNIV ) UNIV KANSAS STATE RES FOUND.
Db	322 LIGRPIKRTDKKGGEAAGAETRYDYLGRFINQSNLKTIDVQGOLCKVATEDIQ 381	XK	XK
QY	1045 TGHMUTTVEFGUNREIGRKLCGSSGHTDILQWSWLKYOQLANRIVKUNGVIQRTQYS 1104	PT	PT Collmer A, Alfano JR, Cartinhour SW, Schneider DJ, Tang X;
Db	382 AGTGYTIDMEYDQDQBLQKLTQTAQNAQALITDOWAVGQKLTKRDLQAGSPHLPTS 441	XX	XX DR WPI: 2003-679632-64.
QY	1105 YDSRNLNQYKQDGAECPDKYKHSIVTONFTYDIGNITACTHPADGTEHATPKFAN 1164	DR	DR N-PSDB; ADL12167.
Db	442 YDPRARLTINLYLSSLSPRLQERLREMTRQFSDELDNLTLCQTRFTDTSERAFKYG 501	XK	
QY	1165 P-----TDPQCQTEVHHTHDPMDPNTRLKDKAGRAVINITDNGENTENFTYDYGRL--- 1216	PS	PS Claim 15; SEQ ID NO 106; 284PP; English.
Db	502 PGDDKHKKDRQCOLLSIAYTPPKTPPTFSYDANGN---QLKDEHN--SLHYDSQSRLLQV 557	XK	New nucleic acid molecule, useful for preparing a composition for
QY	1217 -ONGGQ---STYGYDYPANRUVSQKUTDCL---YXRETMVNEYGENMIRLRTGTETI 1271	PT	PT treating cancer.
Db	558 AETGGAPISQYRYDGHQNLQTR-DGNESEILRLFEGHQLSSTQBDQRTQYDILGEQPL 616	XK	The invention relates to novel Pseudomonas Avr and Hop genes, a sequence
QY	1272 AQQ---RASKULLTGTDSQOSVILTSKDQNLQSEAYSAVQHKSTANDASILGNGERA 1328	CC	comprising a hybridization medium that includes 0.9 x saline sodium
Db	617 GQQTQVWDASQETLILLTIDANGSVMGRFQGQPKARYTAYSGERHSEREALSTAGNGEVE 676	CC	citrate (SSC) buffer at a temperature of 42 deg C. The nucleic acid
QY	1329 FVSGVTHLGNQYRSYSDPTLMRFHTPDLSLSPFAGGGINPNVSYCLGDPINNSRSPSCH--- 1383	CC	molecule is useful for preparing a composition for treating cancer. This
Db	737 TGRLRPDPDEGAALPMQOGGGDINGWVGIGVFTWLVGAATIATGTTATGPTVGLI 796	XK	sequence corresponds to one of the proteins of the invention.
QY	1417 -----AAATSTTALAFG---ALSVTSDITSVGAGLEDASPQASSILGIVS--- 1462	SO	Sequence 982 AA;
Db	797 SMTASAAGAAVNSTVSTGCGTALGIVFTWLVGAATIATGTTATGPTVGLI 851	RESULT	7
QY	1463 -----MONGAAGLAESAIKGKGKLA-----THLGFAEDGENALKTSESSR 1505	ID	ADL12168
Db	205 AGTOKFDGLERVTQTKTCDRVEOFNYDAGEMQ-PRSRITAGLDNINFYTRALDQISS 263	DB	822 RKETDAIGCCTIQYDNYKRVIOITLPGTTVNPNEYKAPEPSTDTLTDIYV--NGIS-- 875
QY	145 TEORDALQOSTLFSYDNNSRMVSVSSTLADGSVINSYAPLOSSELATMLLEVHONGTRIV 204	CC	
QY	876 LGQQTFDGLSLRRTOSQDPGGRWAVTYSAGNDQCPSTVTPDGQPHYQXQPBDLAVQ 935	CC	

QY	936 ASN[IT]OOPSYNPVPGALIKAVARGQSLTPVPPSGRLKMENTNDMKMGSYLN-----T	989	XX	(CROY) CAO Y.
Db	264 TAFDETAKEDYDKTSARLIERATNPQGIRTYRVDHNQJLTGETWDL--LGQANETRHOSS	321	PA	(HINK/) HINKLE G. J.
QY	990 LRGLENGYTDL-----TGPLOKISRSRDTGRVTOIKDSSKTTINDDNLRHGSQVTLA	1044	PA	(SLAT/) SLATER S. C.
Db	322 LGSPRIKRTDKKGGEAGAETRKYDYLGRFRINOSNURTTIDYDVGOLCKATEDIQ	381	PA	(CHEN/) CHEN X.
QY	1045 TGHMLITTEPFDGLAREGRKLCDSSGGHTLDIQQSWLKTQQLANRIVKLANGVQRTEOY	1104	PA	(GOLD/) GOLDMAN B. S.
Db	382 AGGCVILIDNEYDQGQELRTQTASNOALTQ-TOTWADGLKTRDQAGSPLHEFTS	441	XX	New recombinant DNA construct comprising a promoter positioned to provide
QY	1105 YDSNRNLNOYKCDGACCPDQKYGHSIVTONFTDLYGNITACHTFADGTBDAUTFKPAN	1164	PT	for expression of a polynucleotide encoding a polypeptide from a
Db	442 YDPGRGLTVVNYLOSSLPFDELQREMTRIOFSFDELDNTLCTQTRFGTSPERAKYGS	501	PT	microbial source, useful for producing plants with improved properties.
QY	1115 P-----TQCQLTEVHHTHPDMPNTRKQDGKAGRVINITDNEIGENTENFTYPTVIGR-----	1216	XX	Claim 1; SEQ ID NO 11243; 122pp; English.
Db	502 PGDDKHDKRCOLQSLIAAYTPPKTPDPTFSYDANGN--OLKDEHGN--SLHYDOSRLLQV	557	CC	The invention relates to a recombinant DNA construct comprising a
QY	1217 -QNCQG---SVYGDPLNRLVSKQDITDCLB--YRETMVLNEVRNGEMIRLRTGETII	1271	CC	promoter function in a plant cell, where the promoter is positioned to
Db	558 AEGGAP-SQYRUDGHNQVLA-----DGNESL-----STQEDORTQVHNGRPL	616	CC	provide for expression of a polynucleotide encoding a polypeptide from a
QY	1272 AQQ---RASKVLTGTDQSOSVILTSKDKNLSOBAYSAKGKSTANDASINGYGERAD	1328	CC	microbial source. The invention also relates to a transformed plant
Db	617 GQGIUVDABQTLILTDANQSVMGEFOCQLRAKAVISAYGERHESEALLSTAQFNGEVRE	676	CC	comprising the recombinant DNA construct and a method of producing a
QY	1329 PVSQVTHIGNGYNSYDPTMRKPDSTSPEFGAGGINYNSCIGDPTNRSDSGH-----	1383	CC	transformed plant having an improved property. The plant is a crop plant
Db	677 AANGWYLLGNGYRVAYNPLMRPHSPDFLSPABGGVNFTYCGNPIALRDPTGHDASQG	736	CC	such as maize or soybean. The method of producing a transformed plant
QY	1384 -----LSWQ-----AWTGIGMIA---GULLTAGGMAIAAGGI-----	1416	CC	having an improved property comprises transforming a plant with the
Db	737 TGRURRPEDEGAQPMQGSGDIMGVNGVGVWFLVATATLGTATPVGTPVUTGI	796	CC	recombinant DNA construct and growing the transformed plant, where the
QY	1417 -----ANAAIASSTTALAFG---ALSVTSDTSISGALEDASPKAASSIGWVS---	1462	CC	polynucleotide or polypeptide is useful for improving plant properties.
Db	797 SMTASAANAVSTVSTGALIVGTAATSTANTVAIVNN---DQTAEGEVGSGWLGIAAV	851	CC	The recombinant DNA construct is useful for producing plants with
QY	1463 ---MGMRAGLAEBSAIKGTLA-----THLGAFADEGENALLISTSESR	1505	CC	improved plant properties, e.g. improved cold, heat or drought tolerance,
Db	852 PVGLVGFQGAVVARAVAAKVAANAGTIGVRSVSRIG-LIAGGARTISASSAR	909	CC	tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
<hr/>				
RESULT 8				
ID	ADS22210	ADS22210 standard; protein; 1317 AA.	Query Match	5.8%; Score 517.5; DB 8; Length 1317;
XX	ADS22210;		Best Local Similarity	21.2%; Pred. No. 2e-26;
XX	DT 02-DEC-2004 (first entry)		Matches	334; Conservative 194; Mismatches 505; Indels 541; Gaps 72;
DB	DE Bacterial polypeptide #11243.		Query	Match
KW	Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polypeptide.		QY	5.8%; Score 517.5; DB 8; Length 1317;
OS	Bacteria.		DB	5.8%; Score 517.5; DB 8; Length 1317;
PN	US2003233675-A1.		QY	5.8%; Score 517.5; DB 8; Length 1317;
XX	18-DEC-2003.		DB	5.8%; Score 517.5; DB 8; Length 1317;
XX	20-FEB-2003; 2003US-00369493.		QY	5.8%; Score 517.5; DB 8; Length 1317;
PR	21-FEB-2002; 2002US-0360039P.		DB	5.8%; Score 517.5; DB 8; Length 1317;
<hr/>				
QY	372 TTYNMYHLLTSECKQQNYIQTETAVVAAIIGHNFDQSOFQLPKTKETWRSADNSYRS	431	XX	372 TTYNMYHLLTSECKQQNYIQTETAVVAAIIGHNFDQSOFQLPKTKETWRSADNSYRS

Db 294 -----TTED----- PF GKNKTYSPDEHW-- 313  
 Qy 432 EITET-----TFDESGNPLTKVKIKKKTOKKISPSSTHWEYYPAGEVNDCPPEPYG 482  
 Db 314 ELNETNQLYTISAYADENGNRISVTNENSKTKLA-----YDANGNIK-TTNPLG 364  
 Qy 483 FTREVKKLQTPTDSEPKDPKFIQYRSLQSOSHYTLKEBHRYSATOLANS----- 537  
 Db 365 YSK-----SMTDS-----KNNLISOTDELAKHTSPEYDDNSNLKISDALG 406  
 Qy 538 --TLFQYQWNTDKSELGRLLKQTECTK-----GENGK-----TYSWHKFT 574  
 Qy 407 HETVFSYD-----KYQVIGETBSNKATPSYNNQGDQITITDANGKSAFTYDTVGRT 462  
 Db 575 -----YTKQDDTLOOSHSTITHDNFTHRSQVRSRYTGRL-----FSDTDKDI 618  
 Qy 463 TKAQAKGRYTFQDALDNLSITD-----PMQQTTSNTMPLIEKI-VLQMIKVDXSI 516  
 Qy 619 VTMQSYDKUGRLJTRTLNSGT-----PYA-----NLTLYDVEALNQUDDRPFPVIT 665  
 Db 517 LTTIXKSLXJKRMNQWEXLGTNMMLALWPROTQGHKTSYDPLNRQ-----VS 568  
 Qy 666 TTDVNGNQLRNEFQDGAGREVS-----OCLKOSDGDKFYTIHTQQDE 708  
 Qy 569 VTNALKGKTRNQYDAIGNKISITNAYGKSTRYSNSLMQVLVKUTNAMGK--VVR NYDA 625  
 Qy 709 QGRHHTSTYSD-----YITNGRQOTDPDKVHLSMSKSYDNNQOIANTHWSYGSEKIT 761  
 Db 625 VGNLISTDENGHKINYGDSLNLQSVUVDALRKTRNQYDAYGNKISITNAVKGKSTRYS 685  
 Qy 762 VDPITLTATRQLOQNSNNQYTGKFTYTPSQOPQIQTILEDAEHLQSCHTLTDGWDRV 821  
 Db 686 Y-----NSLN-----OLVKVTDAMGSVVRNTYDAVONLIS----- 715  
 Qy 822 RKETDIAQGCTIQYQDNYRVRIOITLPDGTIVNRYKAPFSTDULITDTRVNGISLGQOPT 881  
 Db 716 --TWDANGRKTNYGYDSLNRQSVTINALGKTRNQYDAVENKLSSSTDANWR--LTKSY 770  
 Qy 882 DGLSLRULTOSDG-GRVWATVYSA\_GNDOCPTSTVITPD--GQFIMHQY-----OPELDDAV 932  
 Db 771 DSLNLVLUVKYTDAMGGCVVRITYDAYGN-----LISTDAKGKHKTUDGYSLDRQVSTDPL 825  
 Qy 933 LOVASENEITQOFSPNVPVTGALKAKVAEGOSLITPTYPSGLKOMENINDKRMMSLWTLRG 992  
 Db 826 GRTARANK-----YDAVGNKISSTDSDGKITSYGVNLRLTKSYIPDQKVSY----- 873  
 Qy 993 LENGTHDLTGTQIKISRDTHGRTQVTOIKDSIKTTUNYDILNRT-----GSOV--TDLA 1044  
 Db 874 -----NYDAVGNRLLTMKDSHGTTAYKDYKDLNRLJSVLNPFDGQKVSYTK 918  
 Qy 1045 TGHMU-----TTVEFGNLRRBIRGKLCDSSSHSTLD--IQCWSLWKTQQLANRIVKL 1093  
 Db 919 VGNRKTOMTVPGDKTTSYSDAUNRIG--VIDSBORHITSYSYAKGNLKTUMTENGV-- 973  
 Qy 1094 NGVLTORTEQYSDSNRNLNQYKCDAECPTDLYKHSIVTOMNFTNLYIGNITACHTFADG 1153  
 Qy 974 -----KTE-YSDKANRL-----VELINKNTQVSSSYKTYLDAAGRNRLKUDQLABG 12020  
 Qy 1154 TED-----HATEKFANPTDPCQLINEVHTRPHDMEDN--IRLKDKAGRVINIT- 1199  
 Db 1021 VESGIBSELKESOLLTMTYGY--DRKLVRKLV-----DYSNKTYSKUDMGMRISMTR 10722  
 Qy 1200 --DNHONTENTYDFTGRLONGQGSVYGYPLNRJQSOKDUDCBLY YRETMVNEIR 1256  
 Db 1073 NVDGIGSTISVYKDADQLOSNTSYSVSKGNLKKRNSTQMSYDE--ANRIK 11229  
 Qy 1257 N-----GEMTRLLRT-----GEMTAQORASKVLTETDQSVIL 1293  
 Db 1130 NVSEFVSNNTPKSSYNFETDGDGMRRIKTINGE--NAQSTKVLDSINSALEPQVLTE 1185  
 Qy 1294 SDKQNL-----SQA-----VSAVGHKSTANDASTL----- 1320

Db	1186	SDTRKNTCTVYGTDLISMTNSENAFTYYWDGLGSRSLSLSDSKKIIKVLYDAGQQVK	1245
Qy	1321	-----GYNGERADPVGTVTHLNGNGYRSYDPTLMLRHTPPDSL-----PFGAGGINPVSY	1369
		:     :   :   :   :   :   :   :   :   :   :   :	
Db	1246	EIGTVDNEFRFTGQEOMDDEGLIVRLARY--YDPGVGRFPTKDVIKGRRTQSiNRYV	1303
Qy	1370	CLGPBINRDRPSCH	1383
		:   :   :   :	
Db	1304	TNNNPVNLDLTGY	1317
RESULT 9			
ID	ABG31849	standard; protein; 2334 AA.	
XX			
AC	ABG31849;		
XX			
DT	05-NOV-2002	(first entry)	
XX			
DB	Human kinase, MEKL.		
XX			
KW	Extracellular signal regulated kinase; hyperalgesia; surgery;		
KW	opioid withdrawal; pain sensitisation; analgesic; chronic pain; ERK;		
KW	MEKL; human; enzyme.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200258687-A2.		
XX			
PD	01-AUG-2002.		
XX			
PP	25-JAN-2002; 2002W0-US002128.		
XX			
PR	25-JAN-2001; 2001US-0264336P.		
XX			
PA	(TEXA ) UNIV TEXAS SYSTEM.		
XX			
PI	Gutstein HB;		
XX			
DR	WPI; 2002-608414/65.		
DR	N-Pшиб; ABK90804.		
Reducing or reversing tolerance, physical dependence, hyperalgesia, for withdrawal symptoms, or pain sensitization in patients on analgesics for chronic pain, comprises inhibition of the extracellular signal-regulated PT kinase (ERK).			
PS	Disclosure: Page 156-161; 163pp; English.		
XX			
CC	The invention relates to a method of reducing or reversing tolerance, reducing the risk of physical dependence or hyperalgesia, reducing the symptoms of opioid withdrawal or inhibiting pain sensitisation in a patient taking analgesics. The method comprises administering an analgesic and an extracellular signal-regulated kinase (ERK) inhibitor comprised in a formulation to reduce or reverse tolerance, risk of physical dependence, hyperalgesia, symptoms of opioid withdrawal, or inhibiting pain sensitisation in patients taking analgesics for chronic pain or those undergoing surgery. The present sequence represents the amino acid sequence of human MEKL (not defined)		
SQ	Sequence 2334 AA;		

QY	146 TGFNNNAFDLKVKPKLL-----NPGAHAIYDNN-----FEATOPRLNRIVDDGH-- 192	Db	1055 ---QTNAYFNKKGKLQLQKVQDGHNNAVTVYNDQNLTAITDASGRKLTFYD-B-NGHFT 1060
QY	193 -----DIPINLEY---QGII-----KTTILPFG 214	Db	1061 SITGPNNKKCVTYSVENDLKKVVTIDGTWTSYDSEGRLVKQVSANSTEAKPVFTEY-- 1118
QY	215 QKEGYRTERFLNRLQNLSTHNF-----LGNNPLTMWSFGYTPPIKGNGILQWMTNT 267	Db	1119 QYSGHRLKGIAKACKEVTVSYDADKKTLMTOBNGRKVQYMEAGNP-----IQVJ 1172
QY	268 APGLAKETVN--YNNN-----OGHHFPOSANLPVLPVTLMKQNPAG 309	Db	1173 DAEGLKITNTKCEGGNNVUTDVDPNDVSGKATESYQVQDGN-----VTSVDAEGT- 1225
QY	310 OPATQAEVSYTSHVYVGCGSNGIWNWNKLNLYGMLTEVNGSTESTRRYDKEGHQIVRI 369	Db	1226 -----ETEYINKNNDV-----TRKDTEGVNTDIAVGLDVASETDOSGKSSAV 1271
QY	370 ERTVNYHULTSBCKQQNQYIOTHTAVAILGNFDSQPSQFLPKTKTETRSADNSY 429	Db	1272 YDKYGNQLOSSKDISASTNWK-----DGSFEADESGMNTASKO----- 1311
QY	430 RSEYETTFDESGNPLTKIKDKKQKISPS-----THWEYYPAGEVDNCPPPGFTR 485	Db	1312 RRKLS--VIADKSG----VLSGSKALEVLISQSTSAGTDGYSATQTELEPNNTYLSG 1365
QY	486 FVKKII--QTPYNSBFKODPEKTYQY--RYSIJLGSOSHVTLKIEERIYSATOLINS-- 537	Db	1366 KIKTDLAKSRAVFRIDLRLQDKRQIOMWHNBALAGKNDWT--KROITFTPANACK 1421
QY	538 --TLPQYNTDKS1GRL-LKQTCCTKGNGKTVVHPTYKODDTQQS-HSITTHD 592	Db	1422 AVVMEVDRHKDKGKGKADEVQELEKGESSSNPQNSSFSTATEENNVSGASVDBE 1481
QY	593 NFTIHSQVRSVYRGVFLFSDTDTKDIYQMSYDQKGRLLRT-----L 635	Db	1482 GFNDUDSLKARTSASQAGSVTKTUVLGSANDKPKVYLTLTSKSKASSVKFTDEKDYL 1541
QY	636 NSGTPPYANTITYDVEI--NNLQDNRPPFVI----- 664	Db	1542 QANVYTYADGSTGIVYNAFKFFSGTQBWNRAAVIPKPKPINKVUDISILQKSATGTVWFDDI 1601
QY	665 -----TTDQNGNLRNETDAGHRHVSOCIKDSPGDKKTYTHTQOYBQGRHET 714	Db	1602 RLIEGSSLTKYSTYDSSNGNTVKERDEBLGATS--TDYBTGK--KTSETDKGEKIT 1654
QY	715 STY--SDYLTNGRCQDPDKVHLSMSKSYDNWG-QIANT-----HWSYGSEKI-- 760	Db	1655 YTYDQADQTLNTMLNTGNTSILH----SYDKEGNEVSKTRAGADQTYKEYDVGKLUV 1709
QY	761 TWDPLTMTKQLOSNNSNNQQT----GKEUTTYFBSSQPIQITFDEAGHLQSCHTLIRD 816	Db	1710 TTDBLGNVLASEYVPLANSNLTKTSPNGEV----- 1744
QY	817 GWDVRKRETAIGCQTIVQDYNVRVIQITLPGTIVNRYKAPRSTDTLTDIRUNGISL 876	Db	1745 GTDRVKSKSYINGTEKIFTYDNGN-----EFSVVNEQN-----TT 1781
QY	877 GOOPFDGLSLRTOSQDGGRWVAYTWSAGNDQCPSTWITPDQFTHYQVQPELDDAVLQVA 936	Db	1782 KKRFDNDKNRLLTELDRGCSQWTWPSDSDKLKT-----SMH-----G 1821
QY	937 SNEITQFSYNPVTCALLKAVAEQGSLITPIYPPGGRKMEMINMDMKKMSYLTIRGLENG 996	Db	1822 DOKEMNOFTN-----KUDOMEMKDPSYSFYDEN- 1854
QY	997 YTDLIGTIQKISRDTHGRVQIKOSSIKTILNYDDLNRHGSQVTDATGMLTTFED 1056	Db	1855 -----GNVQ-----TPITAGGGTSFSYVERNLVSLHKGDKGDLIESVEY- 1998
QY	1057 GLNRBIRGKLCDSGGHTLDIQSQWLKTOOLANRIVKLNGVLTQREQSYSTDSSRWRBLNQYC 1116	Db	1899 -----DANGRNITNS-----ASGKV-----QYEGKLNQV-- 1927
QY	1117 DAGBCPTDKYGHSTVTONITYDIGNITACHTFADGET--HATPKFANPTDCPCQJTEV 1174	Db	1928 -----ETHDGTVI--EVTYDGGNKRV-TTICKGEDKPFATTYKIDEGRNRIKTVN-G 2032
QY	1175 H-----HHP--DMPDN-----RLKDYKAGRVINITDHG 1203	Db	1974 NDESISYDKNGNRTSDGKFTYTWDAEDNITAIVKKGEDKPFATTYKIDEGRNRIKTVN-G 2032
QY	1204 NTENFTYDITGRLQNGQSSVYGYQDPLNRLVSQKUDTLOCELY-YTREMVLNEVENGMTR 1262	Db	2033 KVTRYFYG-----DSLNLYVETDADNNVNTSYTG-----SGQLS 2070
QY	1263 LRITCETTIAQORASKVLITGTDQSOSVILTSKDLKNSOEAISAYGK-HKSTANDA--S 1318	Db	2071 YTENGKVKYFHYNAHGDTIASSTGKTV-----AKYQYDAGWNPTKTEADEVKN 2122
QY	1319 LGTNGERADPVSCTVHGLNGYRSYDPTLMRFHT--POSLSPFAGGGINPPSYCILGDPN 1376	Db	2123 RYRYAGYQDFTGELLYLMARY--YEPRNGVFLSLDPDGSDGSDLSLDQNGYAYGNINPM 2180
QY	1377 RSDSGHLSWQAWGIGMIGTAGLILLTAVGGMALAAAG---GIAAATISTTALAFGA 1432	Db	2181 NVDPGHH-W-WVWLVVNAGFA-----AYDGKAYKSGKGWKGAWAAAASNFGPKLFG 2231
QY	1433 LSVTSDTISVSGALEDASPKAASSILGWISMGM-----GAAGLAESAIKGTTKLAHLGAP 1488	Db	2232 ASRAYFKT-----KAVKITCHGRHGLNQISGRNG-----GRGVNLREKLN- 2273
QY	1489 AEDGENALIKSTSESSRIKGWVTSLSIRE--IERNBEGQVTKDHSRGYTDNFMGKGEQAI 1546	Db	2274 -----VRSPKKVQKONGATKVUGKCATVULKRGKVITAVG-----SS 2312
QY	1547 LVHGDKDGTLYHTGK 1563	Db	2313 RAKGSKHVFTHGKGNK 2329
RESULT 10			
ID	ABU18641	XX	ABU18641 standard; protein; 2234 AA.
AC	ABU18641;	XX	
DT	19-JUN-2003 (first entry)	XX	
DR	Protein encoded by Prokaryotic essential gene #4168.	XX	
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.	XX	
OS	Bacillus anthracis.	XX	
XX	WO200277183-A2.	XX	
PD	03-OCT-2002.	XX	
PP	21-MAR-2002; 2002WO-US009107.	XX	
PR	21-MAR-2001; 2001US-00815242.	XX	
PR	06-SEP-2001; 2001US-004893.	XX	
PR	25-OCT-2001; 2001US-0342923P.	XX	
PR	08-FEB-2002; 2002US-00072851.	XX	
PR	06-MAR-2002; 2002US-0342699P.	XX	
PA	(ELIT-) ELITRA PHARM INC.	XX	
PJ	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen Kl, Zyiskind JW;	XX	
PJ	Wall D., Trawick JD, Carr Gu, Yamamoto R, Forsyth RA, Xu HH;	XX	
PN	INT' 2002 22002/000	XX	

DR N-PSDB; ACA22511.

XX PT New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 46565; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WOPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 2234 AA;

Query Match 4.3%; Score 386; DB 6; Length 2234;  
Best Local Similarity 18.7%; Pid: No. 8.3e-17; Matches 339; Conservative 236; Mismatches 633; Indels 608; Gaps 76;

Qy 21 FTOQANMFTSAVSGGVDPR-----TGLYNNQITLGHIVG-----NGNL----- 57  
Db 748 YSKVGHASEISHEPVPRFSDSGFGFMFDYTWASIVPVLVGRKVNATGNFIMSEKDITLSC 807  
Qy 58 -GPTLPLTLSVSPNKTDIGGIGGFNFGLSV--YDKNSLISLS-----TGEN- 102  
Db 808 RGPDVSVERTTNSQSKEVKGUGLFGTGMSGLERVRWADGNGNLLISTDGANTITPRTGDNK 867  
Qy 103 -----YKVETDKVVKLOOKKLDNLRFEDKLKENCYRIHKSODIEBV 145  
Db 868 YQAPTGIVYLEIKQVSGGYEIKDKDQTV-----TFYKSGDAQR 905  
Qy 146 TGFNNNAFDLVK-----KKLNPAHGAIYIDWNPEATOPRINRIVDPLDHDHDIPL 196  
Db 906 IBYTKOKYGNMTTYEDGASRSLKVKNSAGKGLVL--QYDNNKKCARVI----GPDNKT 959  
Qy 197 LALKEYQQLIKIYLTPGQ--KEGIFTELRLRLOAISIHNIFSLGNENPLTWFSFGY---- 250  
Db 960 ITPNVYGDLLSSTTPBGKVVKYGV-----NGVLTSPVQHDTDAKPTKTSVAYENDR 1013  
Qy 251 -----TPIGKQNL---GQWITSMWTPGGKJKEVWVSNRNGHHRPOSANLPVAVYTL 301  
Db 1014 LUKVTPDPLGKATTALNTGSKEVLTNPKGKIVVYND----AGNPV----- 1057  
Qy 302 MKQVPPAGOPAQAEVSYSITSNYVGG-----GSNIMWNKLDMYXGLMTBEY- 348  
Db 1058 -KTVEDVGRNLNTSYBNAINLKVTTPKQNETATYDGIGNVTSDENGTEKEVYNK 1116

OY 349 -YESTESTRSRYKODE-----GHDQIVRERTVNYHILTSBECKQONQSYIQTTETAYAI 400  
Db 1117 DNGIKIATDNEDRKTTVAVGANTEVSQTDQGAN----TSSVIIHDOQYGNPTEKSL 1171  
Qy 401 IGRNFDSPSOSOFOLPKTKETWMSADNSYRSBETETTDEBSGNPLTKVTKDKKTIKISP 460  
Db 1172 AGGNLIONPS-FEM--NGSTEKKWKVUDTNNSGS-----KDATPAPGGGGESLKITK 1223  
CC 461 STH--WEVYPAGEVDUNPPEPYGFRTRVKKI----- 503  
Db 1224 ATNDWDGYTAATORVLTENTTITLSGMVKTDLVNGAAFFNVQSLNEANGIBOGWHMR 1283  
Qy 504 EKFQYRSLIGSOSHVTUKEHRSATQMLNSTLQYNTDKSELGL-LKQTECTGE 562  
Db 1284 HNKVQGTSUWNRQ----VTPKTTSQ----TRKVIYQLOVENG-----QTPVSEPKDP 1337  
Qy 563 NGKTYSVVHKPFYTQO-DDTLQQ-SHATTH-----DNFTIHSRQSRVYTRGRFS 611  
Db 1338 VSSSFNPVLIASSFBENWPDPGVPOWRSCSQHCRNDVSDSDSTFGHSSIVMER----S 1391  
CC 612 DTDPKDIVQTMYSVTKLGRULTRTNSCPYANTLTYDTELNNLDDNNPPPVITTDNG 671  
Qy 1392 EYEGNDI-----GYNRVRVNLNOKAETVTUTAMRSERNVND----- 1428  
Db 672 NQLRNEPFDGAGRHSVQCLKDSDGDKFTHTQOYDCEGRHHTSYSLTNGRQQTIP- 730  
Db 1429 -----APDKLUSKDYAVLAETTYQG----TWNVNTYSPSGTND 1464  
Qy 731 -----DKVHLSMSKSDVNWQIANTHWSYSVSBKTVDPTPLTAKQQLQNS 777  
Db 1465 NRSAAVIAPKPIKIEFLFLFRKNG-----KWFDDPLLEGNALINQ 1511  
Qy 778 NNVOTGKETVTPSQQPQITLDEAHLQSCHLTDGWDVRKETDIAQGCTYQD 837  
Db 1512 YD-NDGNNVATY-----DEBEGOK--NTFTDAGNPKSEDEKGNTKLYDYN 1555  
Qy 838 NYNRVIQIOTLPDGTVNRYKAVPSTDTLTDIRVNGISLGQQTPDGLRLTQSODGGRW 897  
Db 1556 KDNLITKVKTLNGTSVNRYD----- 1576  
Qy 898 AYTISAGNDQCPSTVITPQDFQFIHYQYPELDDAVLQVASNETTQQFSYNPVTPGALLKAV 957  
Db 1577 ---HNGNTTEKSUMFGKIQTHKVEYDNDKNTVYDALNRREN-TVDENANKIK----- 1631  
Qy 958 AEGOSLITPTVYPSPGRLKMININDMKMMSYLTWLGRGLENYTDLGTICKISRDTHGRVTO 1017  
Db 1632 PNGSILSVEYDADRVVGEK-----RNGKOSFT----FERDONGQVTK 1670  
Qy 1018 IKD--SSIKTTLNTDLMNHIGSOVTPLATGHMLTTVEDGLNREIGKLCDSGGHTRD 1075  
Db 1671 VKLVNGVERTKTKDADR-----VTS AT-----DSRGKID 1702  
Qy 1076 IQSML-----KTOQLANRIVKLANGVQORTEQSYDSRNRLAQ--YKCGAECPTDKY 1126  
Db 1703 ---WAYDKANSKTEKQBLQVQGYY--TNKWSYD-YNTLQDQNTVTDGSO----- 1748  
Qy 1127 GHSTVTONFTYDINGNTIACTHTPADGETDHATRKF- ANP-----TDPCLQTEVHNT 1177  
Db 1749 -----TFRFDDQAVV-----RTTAGNGSSGSTRYDQANKIKDOLVUGTSNSILSERVE 1799  
Qy 1178 HPDMDNIRLKQYDRG-----RVINITDNGNTENFTYDITGR-----LQ 1217  
Db 1800 YDQSGNRTKIKHEGAGGKVETENFWYDIPINQOLNEVLPNPGNTKSYTYDGFGNRTSVKVIC 1859  
Qy 1218 NGQGS-----VYGDPLNRLYSQKUTIDCSL----- 1244  
Db 1860 NGKETKSIATMENSGNQVKPGNESLTYDVGNGRTSDGKVKYTNWEDDQIVAITKOGENN 1919  
Qy 1245 -----YRFTMLVNEVRGEMIRLRLRSETI-----IAQO 1274  
Db 1920 AFATVKDEDNRFEKVNQCVTFYFDDGSINLYEVTOGNTLVRQVYVSDGARLAMK 1979

1275 RASKLUTGTDQDQSVI--LTSKDKNLQSQEYASIGSK----HKSTANDASILGNGERA 1327

1980 AQQGTYHNPGRGVWAMTNQKVETAVYDANVNLTSDTKGAAAD-NPFGVAGYMY 2038

Db 1328 DPVSQTHLGNGYRSYDPTLQRFHFD-----SLSRGAG----GTPNPSYCLGDPINR 1377

Qy 2039 DKEIGMYL-----IARYVPERHGFLVSDDPGDEDPPVTMNGYVYADNNPVM 2088

Qy 1378 SDPSCHLSWQAWTGIGMGIAGLLTATGMAIAANGGIAAIASTSTALAFGLSVTS 1437

Db 2089 TDPOCKMW-----LVPVVIAGAMVAAPGAKYAI---RYGAKYKKAUKS 2131

Qy 1438 DITSVSGALEDASKASSIGLGMWSNGMGAAGLAESAIKGKTKLATHLGAFABDGEMALL 1497

Db 2132 -----GWDY---GKRVAKSGWNKKSIAQKIPRITHKGR---I 2163

Qy 1498 KSTSSSRKMGVTHSLDREIVRNBERGQVKDHSRGYTDNFMKGKBOAQLVHGDKGFLY 1557

Db 2164 KGDNOKKGWGVWITISTK-----KIGKRIVS-----SFERTPPHNGHY 2203

Qy 1558 HTEGKIKHNG-KGPYTR 1572

Db 2204 HLQKMKYSKYQGKWR 2219

RESULT 11.

ADC01365 standard; protein; 1400 AA.

XX ADC01365;

AC 04-DEC-2003 (first entry)

DE Enterohemorrhagic E. coli 0157:H7-specific protein SEQ ID NO: 1410.

KW enterohemorrhagic; anti-bacterial.

OS Escherichia coli; 0157:H7.

XX JP2002355074-A.

XX PD 10-DEC-2002.

XX PR 24-JAN-2002; 2002JP-00015959.

XX PR 24-JAN-2001; 2001JP-00112010.

XX PA (UETS-) UNIV TSUKUBA.

XX DR WPI; 2003-451640/43.

XX PT Enterohemorrhagic Escherichia coli 0157:H7-specific nucleic acid molecule and a polypeptide and its use, a polypeptide, a vector and a host cell.

XX PS Claim 3; SEQ ID NO 1410; 2067pp; Japanese.

XX Qy The invention relates to a novel enterohaemorrhagic Escherichia coli 0157:H7-specific nucleic acid molecule. A polynucleotide of the invention has anti-bacterial activity. The polypeptide can be used in detection and/or treatment of 0157:H7 infection. The nucleotide sequence of the genome of Enterohemorrhagic E. coli 0157:H7 was determined. The present sequence represents an E. coli 0157:H7-specific polypeptide of the invention.

XX SQ Sequence 1400 AA;

Query Match 41%; Score 365; DB 7; Length 1400;  
Best Local Similarity 20.9%; Pred. No. 1.2e-15;  
Matches 284; Conservative 154; Mismatches 467; Indels 456; Gaps 65;

Qy 418 KTETM-RSADNSYRSRSEITTFDESQNPPLKVI-----KKTTQKISPSITHWEY 466

Db 218 RTLTTRREAGDLAGLIGTGT-DAGEKFPRMLITIQQARBARAKHTTASISPPD---- 272

Qy 467 YPPAGEVDNCOPPEPYGFTRFKKLIQTPDSEFKDPEKETIQYRVSLSIGQSHVTLKIE 526

Db 273 -----PRELSDSAPDPTIPTGTEY----GPDGIR----- 297

Qy 527 RHYSATOLNSTFQYNTDKESELGRLLKQETCTKGENGKYSV-----VHKFTYTKO 579

Db 298 -LSAWWLJHDPAVESPGLGAPLARY----TYEAGELLAVIDRSNTVRAFTYDAH 349

Qy 580 DTLOQSHSITTHDNFTIHRQSRSRTGFLFSDFDTDKD1V7QMSYDKLGRLTRIANSR 639

Db 350 PGIMVAH-----RYAGR-----PEMRYVYDPTERWVQEINPAGI 383

Qy 640 PYANTLTIDYELANNLQDDNRRPFPVTTD-VNGMQLRNEFDRGHSVQDLSKDSGDGK 698

Db 384 SY----RYQWE----QDR-----ITVTDISLNREVHLTEGGAG--LKRVVKKELADG-- 425

Qy 699 YTHTHQDQVDEQGHHTSTSYDYLNGR--CQTDPDKVHLSMSKSYDNWSGQIANTHWSGV 756

Db 426 -----SVTHSGYDAGRLTAQTD-----AAGRTEYGL 453

Qy 757 SEKITVDPILTATKQLOGSNNNNVQTGEVTTYPSQQPQIQTLEDEAGLQOS---CHT 812

Db 454 N--WVSGDIPDITPDPGRBTKFVYNDGNQLTAVSPDGLSERRAYDEPGLVSETSRCGO 511

Qy 813 LTRUDGWDRYKE---TDAIGOCITYDYNVNVIQITLUPDTGIVNPKYAPFSTDTLT 867

Db 512 VIRAYDNPHSELPAFTTDTGSTRQMTWSSRGYQALFTDCSGIQTRYBDRFGOMTAH 571

Qy 868 DIRVNGISLGQQTPDGLSLTQSODG--GRWAVWYMSAGNDQCPSTVTPGQPPIHYQYQ 926

Db 572 --RREGIS-RYRDRDNRCRLTSVDAQGHTERYTNAAGDL--TAVITPBNRSETOY-- 624

Db 927 ELDDA--VLUQASNBEITQOFPSNTPVTGALLKAVAGQSQSITPIVPSGRKLKENINDMK 983

Db 625 --DAWGKAVSTVQGGLTRSMEVYD-LAGRTITLNENGSSRFETYDA----- 667

Qy 984 MSVYNTLQGLENGTYDFTQKISRDTRGVRTOIKDSSIKTTANDDRNHIG----- 1037

Db 668 LDRLVQQRGFD-----GRTQRYHDLTGTQKLTQSEDEGLVTLWHYDESDRLTHRTVNG 720

Qy 1038 -----SQVTLAIGMLITTEFDGLNREIGRKLCQSSGHTDI----QGS 1079

Db 721 PARQWQYDHRGMWLEFISHLSEGHQAVAVRYGYDDKGKLAGCRQTVNPERGELLWQHETEH 780

Qy 1080 WLKIQOQLANRIRVKLNGVJORTEQSYDS-----RNRL-----N 1112

Db 781 AYNEQGLANRVTP--DSLPRVEMLYTGSCYLAGMKLGGTPLVETDRDHLARETWSFGNN 838

Qy 1113 QYKGDGACCPPTDKYGH-----STV-----TONFTYDYGNT 1144

Db 839 AYELTSTVTPA--GHLSQQRQLSAQVYDYYDWDNGDVLWRSQPROTWYGSANGRL 895

Qy 1145 ACHTFADGTEHATFKFAMPDPC--QLTVEHHTPD----MDPNIR-----KX 1189

Db 896 SVRTLSD---LDIPIYATDAGNRPD-PEIHPDSLTAMPDN-RIADAYVYRH 948

Qy 1190 DKAGRVINITD-----NHGNTENFTYDTCGL-----ONCGQSV--YGDPL-- 1229

Db 949 DBYGRLTBKDRIPAGVIRTDDERTHHTHDSQHRLVFTTRIOHEPVLVSYRLYDPLGR 1008

Qy 1230 -----NRLVSOQKDTQCELYR----- 1247

Db 1009 RMAKRVWRRERLDLJGWMMSLRKPEVWTWYGMGDGRLLTVQDPTTRIOTVFBGSFTPLRV 1068

Qy 1248 -----ETMLVNEVRG-----EMIRLJRTGETIAQORAS----- 1277

Db 1069 ETENGEREKQRSLSAETIQEQSESGHVFPAELVRIDLREIRADRVSSRAME 1128

Qy 1278 -----KYLTDGTDQ---OSVILTSKDLQSQEYASIGKST 1313

Db 1129 AQCLGTVOLARQVEPETTPARKVHFICHDRLGLPLALISEDGNTIAWRGEDEWGNQNLNE 1188

QY 1314 ANDASI--LGNGERADPVSGVTHLGNQYRSOPTMLRHTDLSLRG--RGINPVSY 1369  
Db 1189 ENPYTLHOPYRLPCQHDEBESGLYNNRY--XDPLOSYRTD--PIGLAGWNLY 1243  
QY 1370 CLGPDPINRSDPSG---HLSWQAWTGICMGIACLLT----IATG---GMAIAAGGI 1416  
Db 1244 PL-NPIIRDPPLGJYNYQOLLDWHDHDDSYGTSSIDITSGDLISLGHGUVAFAKK 1302  
QY 1417 AAIAIASTSTTALARGALSUTSITDSIVSGALEDASPKA--SSTLGWWSMGMGALESA 1474  
Db 1303 GEMLSDICIIYATACCHAGIGGGINAIITYSYSKSLPTGSVNSVG-VTGGVGG---- 1356  
QY 1475 IKGSTKLATHLG-AFAEDGENALIKSTSSESSRIKGWVRL 1514  
Db 1357 -----HPAYTVVWDVNDP--ESSTSVGIGAGVDAV 1386

RESULT 12

ADG30698  
ID ADG30698 Standard; protein; 1510 AA.  
XX  
AC ADG30698;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Xanthomonas axonopodis pv citri plant pathology-related XAC3245 protein.  
XX  
KW Xanthomonas microorganism; plant; pathology; bacterial pest; xac; xcc;  
XAC.  
XX  
OS Xanthomonas axonopodis pv. citri.  
XX  
PN WO2003089647-A1.  
XX  
PD 30-OCT-2003.  
XX  
PR 22-APR-2003; 2003WO-BR000060.  
XX  
PR 22-APR-2002; 2002US-0374620P.  
PA (AMPA-) FUNDACAO AMPARO A PESQUISA DO ESTADO.  
XX  
PI Da Silva ACR, Farah SC, Quaggio RB, Reinach FDC, Ferro JA;  
PI De Oliveira JCF, De Laia ML, Setubal JC, Furlan LR;  
DR WPI: 2003-86544/80.  
DR N-PSDB; ADG30697.  
XX  
PT New nucleic acid molecule from a Xanthomonas microorganism, useful in  
determining the presence of Xanthomonas bacteria in a sample.  
XX  
PS Claim 8; SEQ ID NO 55; 145pp; English.

The invention relates to a novel isolated nucleic acid molecule from a Xanthomonas microorganism where the nucleic acid molecule is associated with pathogenicity caused by the Xanthomonas microorganism, or its variant, that causes reduced or enhanced pathogenicity. The nucleic acid of the invention may be useful in detecting the presence of Xanthomonas bacteria in a sample, as well as in plant pathology. For example, for identifying nucleic acid molecules and proteins involved in pathology caused by bacterial pests. The current sequence is that of the Xanthomonas axonopodis pv. citri (xac) plant pathology-related xac protein of the invention.

SQ Sequence 1510 AA;

Query Match 4.1%; Score 361; DB 7; Length 1510;  
Best Local Similarity 21.2%; Pid: No. 2.5e-15; Mismatches 304; Conservative 158; Mismatches 466; Indels 504; Gaps 75;

QY 144 VLTGFNNNAPFLK-----VPKKUJNPAGHAIYDNNEFATQPRNLRYDDLG 191  
Db 250 IILTRPNCNFYTFKKSANGTWIPDMDVRETLISRVQDNGTLTGWQVATDDSKEQF--DLDG 307

QY 192 HDIPLNLNEYQGLIKITILTFPQQKECYTERFLRFLRQLNSIH-----F 236  
Db 308 ---KLTGJSY-----TQCHQLTLYTASGRQLOQSVDTRGRLLFAYQADRT 350  
QY 237 SLCNENPITWSTCGTYPTRIGKNGLIGQWMTSMTAPGLKETV--NYSNNQGHFPOS-- 290  
Db 351 QVLGPDGVNLAVGQDSDLR---RSVLTQTAGAVASAAIGYDGN---ARFPDALTT 402  
QY 291 -----ANLULPVPTVLMQKVPGAGOPAI-QAEYSY-----TSHNYVGG----GSN 330  
Db 403 HRDQDQYTAWSWTYDAAQRVWVISVHGDPDTGKIDEATTAYSGNISTVSNALGNVAUTRQG 462

QY 331 GIWNWKLDNLIGLMTETWYGSTERR---YKDKEGHQDQVIRIERTTNVHILTSECKQ 385  
Db 523 DNSGGQGRITLQP-----DHHS-FRPV-TDQPTYDASDILVARTSMWTSNRGQA 570  
QY 433 ITEMFDISGNPNTKVKDKTKTOKTISPSHTHWYYPGPGEVDNCPPPEVYFGTRFKKTIQ 492  
Db 602 -----SLRG-----LILASDGRTDIADSTSY-----YL 627  
QY 571 LTKQBTCT-----KOBENGKTYSVVHKFT-YTKQDD---TLOQSHSITHDNFTIH 597  
Db 628 ADDASCTASASTCPHRKDRMKVNAQGQTTELAYDAGRPLSIKDPNGIV-----DYLH 685  
QY 598 -----RQVRSRVTGRLFSDTPTKDTVOMSYDKLGLGULTRTINSGTPYANTLYDYL 651  
Db 686 PRGWL TASKVKG-----ADASSEADRTRIDYWP TG-LVRQVTPQPGAFATFYD-AA 737  
QY 652 NNLODDNRRPPFVITTDNGNQLRNNEFGAGRHVSQCLKDSDCDCGKFVTHQOYDCCR 711  
Db 738 HRLRD-----ITDKNAGNTVHYTLUNAGNVRKVDTDAATLK--RTLSRVNQLGQ 786

QY 712 HHT--STSD-----YLTNGRQOTDPDVHLSMSKSYVNWQI-----ANTHWS 753  
Db 787 LKQDQATAASDPTDFAYDANGNATKUTDIALATQSEVPLNRLSHTLDVAGIKADTHA 846  
QY 754 YGVSBKTT--VDPITLTATKQLOSNNSNNO-----TGKEVTTPSQOCPIOITLFDEAH 806  
Db 847 YDALDNLTKVTDPKGLDITDYNGFGLBLKLTSPDTG--VTST-----YDSGN 894  
QY 807 LQSCHTLIRDGVDRVRKEDTAIGCQTYQDNTNVRVQITLPGTIVNRKYAPFSTDIL 866  
Db 895 RAT-----QFDARGNTTAVASYDALLNRLTKVTPS-----SLDVTY 930

QY 867 T-DIRVNGISLQGQTFDGEJSRLTQSQDSGRWVAYTISAGNDQCPSTVTPDQGFIHYQX 925  
Db 931 TYDVTQACTSG-ETF-SGRLTMQDGAIOTYCOYNFGD-LYRKVQTSNGTALVLRD 987  
QY 926 PEL-----DDAVLQVASNE--TTOQFSYNPVTGALLKAVAEGLSITPVIYPSRL 973  
Db 988 YTVGSQLRRTYDGAIVDVTYRNAQGTTQVGVTPAGES--RQVLGNA--TYPFGER 1042  
QY 974 KMENINDKMKNSVWLTL-----RGLENGYTDLTGICKRISROTHGRVYQIKOS 1021  
Db 1043 AGWTYGNGRITLARQDLDYRPOAQDTPRGEGLIDVGF-----GFPDAGNLITALTPA 1092

QY 1022 SIKIT--LYNDLNRHGSQVTDATGMLTTVEFGUNRERGRKLCDSSGHLIDQ 1078  
Db 1093 GNTTPEIGLGYDAIGRLG--LTDGVUT3WV-----DG----- 1123  
QY 1079 SWLKTOQLANRIVKLNGLVQRTQSYDSR-NRLNQYKCDGAECPTDKYGHSTVQFTY 1137  
Db 1124 -----YSDATGNRL-----STKQG-TTQIQVY 1146

required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at  
 fpi\_wip int/mip/mipshd not sequences

		RESULT 14
ABU19676	ID	ABU19676 standard; protein: 1515 AA.
XX		
AC		ABU19676;
XX		
DT	19-JUN-2003	(first entry)
XX		
DR	Protein encoded by prokaryotic ebsen	
XX		
KW	Antisense; prokaryotic essential gen	
XX		
OS	Borrelia capacia.	
XX		
PN	W0200277183-A2.	
XX		
PD	03-OCT-2002.	
XX		
PP	21-MAR-2002; 2002WO-US009107.	
XX		
PR	21-MAR-2001; 2001US-0081542.	
PR	06-SEP-2001; 2001US-0094893.	
PR	25-OCT-2001; 2001US-0342933P.	
PR	08-FEB-2002; 2002US-00072851.	
PR	06-MAR-2002; 2002US-0362699P.	
XX		
PA	(ELIT-) ELITRA PHARM INC.	
XX		
PI	Wang L, Zamudio C, Malone C, Hassel	
PI	Wall D, Trawick JD, Carr GJ, Yamam	
XX		
DR	WPI; 2003-029926/02.	
DR	N-PADB: AC223546.	
XX		
PT	New antisense nucleic acids, useful	
PT	for homologous nucleic acids require	
PT	isolate candidate molecules for rat-	
XX		
RS	Claim 25; SEQ ID NO 47600; 1765pp; 1	
XX		
CC	The invention relates to an isolatee	
CC	the 6213 antisense sequences given	
CC	of the nucleic acid inhibits prolifer	
CC	(1) a vector comprising a promoter (	
CC	encoding a polypeptide whose express	
CC	nucleic acid, (2) a host cell contain	
CC	polypeptide or its fragment whose ex	
CC	antisense nucleic acid, (4) an antisense	
CC	the polypeptide, (5) producing the p	
CC	proliferation or the activity of a co	
CC	proliferation, (7) identifying a comp	

	Query	Match	Score	Length
Best Local Similarity	3.9%	Score 350;	DB 6;	Length 1515;
Local Similarity	22.1%	Pred. NO. 1.5e-14;		
Conservative	23.2;	Mismatches 314;	Indels 380;	Gaps 59;
Matches	126;			
QY	481 YGFTTRVRKKITIQTYPDSEFKDPEKPIQYRSLIGSQSHVTLKIBERHYSATQLNSTLP	540		
Db	543 FQFGQFYYAVDEHGMWQMRDTQDVRYTRYDTAG-----RVVE-----TGRQ	586		
QY	541 QYNTDKSESELRLKQTECTKGENGKYSTVHFKYTSQSHVTLKIBERHYSATQLNSTLP	596		
Db	587 GY-----GRFTIVEGC-----RVDVGDWEWTYAYNDEGLVTAETDPLGHCTYESEWL	636		
QY	597 HRSQVR-----SRY-TGRFLS-----DPTDKDVT-----QMSY	624		
Db	637 GRMLMARIDPUGRRRTDPRYDERGQLQHSVWESSGRFTWDYDFDDEOLTGARLUPNGSTIKLEY	696		
QY	625 DKLGRLRILTRTLNSGTPYANTLYDY-----BL-----NNLQDDNRPPPVITTDV--	669		
Db	697 DHLISRLIART----BPDGNKNTIYRGPRGBLLRVQGDRETRBLYDR----URLTDEL	748		
QY	670 -NGNORNEPUGAGAHRVSDQCLKDSDGDKFYTINHQYBEGRHHTSTSVDYLNGR---	725		
Db	749 PTGARFRKKIDALGR----LIEETSPFDG----RVTYD----VADGPANPRGLL	790		
QY	726 -QQTDPDKVHLSMSKSYDNNGQIANTHMSVGUSKEITV--DPTLTATKQLOQNSNNVQ	781		
Db	791 SAVTRPD-----GSVSHARYN----SESPVIENTDPLGRTIG-----824			
QY	782 TGKEVTTYTPSQQPQIQLTDFEAGH-----LQSCHTLRDGMDVRK	823		
Db	825 -----TYGPPD----LITASIDAAHGHTREEDYDHATLTKWNINALGETTYTRYDAAGRLAA	876		
QY	824 ETDATCQCTYQDNNRNVIQITDGTINRKYAFPSIDLTDIVNGISLQ-----QTF	681		
Db	877 EIDWGKGRATEYDRAVGRLLKTLPDG-----GOWRTY	910		
QY	882 DOLSLRLTQSODGRVWVAYTISAGNQCPSTVITPQQPIHYQYQPELDLAVLQVNEIT	941		
Db	911 DASDRLLBIDAGDVKAAYRDASG-----RLASAVQGEHTHVT	949		
QY	942 QOFSYNVVTGALLKAVAEQSGSLTPYPSGRKLMENINDMKMOSVYIWTGLLENGYDLT	1001		
Db	950 -RFAYDR-NGRLIGEQHNGELLRHVDAGCQLL-RMTPARETTVAY-----993			
QY	1002 GPTOKISRDTEGGRVQIKODSSIKTTWYDQNLHIGSQVNDL-----ATG-----HMLT	1050		
Db	994 -----DVSGALITYVG----QLTIRRDRGJGREIGREAGDFVAQQOYDALGRIRRQIAG	1041		
QY	1051 TTVEPDSLNLTRGKCLCDSSGHTLDIQQSWLKTQQLANRVIKLNGLQRTQEYQSTDNR	1110		

1111 LNLQYKDCGAECP---DKYGHSTV-----TQNFPIYGNITACHTFADCTEDHATP  
 QY :  
 Db 1073 LERVD-TGADTLTYORDERGOTICAESLQOPSEHFRYDAMNI-AANGORA-----  
 QY 1161 KFANPTDPQCQLTERWHTH---PDMDNIRLKDKAGRGINIT-DNHG---NTENFTYDT  
 Db 1122 ---PVD-----AHYHRGGLBQGVARYKDARGTIKEWQCPVRPKTWOYTWDG  
 QY 1213 LGRLU-----OOGCGSVYGDPLRNLVSQVTPDTLDCLEYRET-----MLVN---EV  
 Db 1172 LNRLVRVVTPERGWAA-YRYDAFNRRIKQ----QGGGRETFKFLIDGPMLAERVIEQ  
 QY 1256 RNGEMIRLRLRTGETI-----INQRASKV- LITGTDNSQSVLTSKDNQNLSE  
 Db 1225 RDG-----TTGOWTWIHEPGSFLPLAQETDDGFLPILTDPQIGRKTVFVBQGRPYWKA  
 QY 1303 AYAYGK---HKSTANDA-----SILGYNERRADPSVGTHLNGQYRSIDPTLMRFH  
 Db 1279 ASLWGKULPVKOPANDAGCTSDITLRFSCQWADDETGLYNLRY- YDPDSQGYL  
 QY 1352 TPSLSPFG-AGCINPVSYCLGSPINRSDPSG 1382  
 Db 1337 SAD---PIGLGGARTQAY-VHDFSQWIDPLG 1364

RESULT 15

ABU15135  
ID ABU15135 standard; protein; 1397 AA.  
XX AC ABU15135;  
XX DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #662.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design  
XX OS Escherichia coli.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PR 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948933.  
PR 25-OCT-2001; 2001US-0342933P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX PA (BLIT-) BLITRA PHARM INC.  
XX PT Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PT Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
DR WPI; 2003-029926/02.  
DR N-PSDB; ACA19005.  
XX PT New antisense nucleic acids, useful for identifying proteins or screens  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX P8 Claim 25; SEQ ID NO 43059; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one  
CC of the 623 antisense sequences given in the specification where expressed  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
encoding a polypeptide whose expression is inhibited by the antisense  
nucleic acid; (2) a host cell containing the vector; (3) an isolated

polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) Profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

Db 115 AERWQDDE-RGMHTDSISSEGHRVTVNGDQKGRASEHUTVHQPTNEILWQHETRH 773  
 Qy 987 LWTLRLGLENGLGYLTGTLTOKISRUDTHGRVTOIKOSSIKNTLNDDLNRRHGSQDLDATG 1046  
 Db 774 AYNAQGLAN--RCIPPSLPAVENLYIG-----SGWLISGMKLG 808  
 Qy 1047 HMLTTTVEP--DGINREGRKLGKODSGHTDLCOSWLKIQQLANTRIVKINGVUORTSOYS 1104  
 Db 809 D-TPLVEYTRDRDLRRETUR----SFGRYELTTAYTPAGOLQSQ--HLSLLS-DRDYT 858  
 Qy 1105 YDSRARNIINOYKCDAECPPDKYGHISIVONFTIDIGNITACTTFAD----GTEPH 1158  
 ::::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 859 WNDNGELIR-----ISSRQ-----TRSYSS-TSTGTRGVHTAANDDIRIPTYTDP 906  
 Qy 1159 TFKTIANPDTDCOLTEVHHTPD----MDPN-----IRLKUDKAGRUVINID- 1200  
 ::::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 907 GNRULPDP-----EHPDSALSMWDNRJARDAHLYRUDRGRETEKUDLIPEGV 956  
 Qy 1201 ---NFGNTENPTYDITLGRILQNGOSSSVY---YDPLNRUVSQKUDTDLCEL--- 1244  
 957 IRTDDERTHRHYDQSQRHLVHYITRQAEPLIVESRYLQPLGRVAKRWRRERDITGM 1016  
 1245 -----YY-----RETMVLNS-----VRUGEMTLLR--TGTITAOA--- 1276  
 ::::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 1017 SLSRKPOVTVWYGDQDRLLTTONRTRIQTYQCSFTPLIRVETATGELAKTORSLAD 1076  
 Qy 1277 -----SKVLTGTQSSQSVLTSQKNTSQEA----- 1303  
 ::::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 1077 TLQQSGGEDCGSVVFPPLVQMLDRLESETLA--DRVSBESRMWLASCGLTVAMQMSOM 1133  
 Qy 1304 -----YSAYGKHKSTANDAS-----LGTING 1244  
 ::::|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 1134 DPVVTTPARKHLHYCDHGRGLPLAISTEGTCAWVAYEDWGNLNBENPHOOLIRPG 1193  
 Qy 1325 ERADPVGSVTHLGNGNQYRSYOPTMLFHTPDPSLSPG-AGGINPVSYCLGPINNSDSPG 1382  
 ::::|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 1194 QQYBESGLY--NKHRYVDPLQGRXITQD--PLGKGQWNPFQYPL-NPISMDPLG 1246

Search completed: January 30, 2006, 09:46:15  
Job time : 140.372 secs



RESULT 2  
US-09-492-709A-340  
; Sequence 340, Application US/09492709A  
; Patent No. 6720139  
; GENERAL INFORMATION:  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Trawick, John  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Froelich, Jamie M.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
; FILE REFERENCE: ELTTRA\_001A  
; CURRENT APPLICATION NUMBER: US/09/492,709A  
; NUMBER OF SEQ ID NOS: 485  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO: 340  
; LENGTH: 1426  
; TYPE: PRT  
; ORGANISM: E. Coli  
; US-09-492-709A-340

Query Match 3.8%; Score 339; DB 2; Length 1426;  
Best Local Similarity 20.2%; Pred. No. 9\_6e-17; Mismatches 409; Indels 458; Gaps 60;  
Matches 257; Conservative 150; MisMatches 409; Indels 458; Gaps 60;

QY 572 KFTTYTKQDITLQSHSITHDNFTTHRQSRSYRFLSDTDKTDIVTOMSYDKLGRUL 631  
Db 317 RYTTEAGEEL---AVDRSRNTORAFTYDAOPGRMVAHYAGRPMRYRDTGKV 372

QY 632 TRTLNSCTIYANTLYDXELANNLDDNNRPPFVITPD----- 668  
Db 373 EQLNPAG----LSYR----LYEQR----ITVTDLSNRREVHTEGGAGLKRVVKKE 418

QY 669 -VNQNQLRNEDFGAGRHSVQCLKDSDGDKFYIHTQWD-----BQGRHHTSYDSYL 721  
Db 419 LADDSVTREGSYDAAGRILAQ---TDAAGRTERGYEUNVSDIDITPDRERETKPYND-- 474

QY 722 TNGRQOT---DPKVHLSMSKSYNWGQJANTHWSYGVSEKITVDPITLTATKOLQNSN 778  
Db 475 --GNQUTAVSPD---GLESERREDEPGRML-----SE 502

QY 779 NVQTCOKETI-TYPSQQPIQITLDEAHLQSCHTLRGWDRVRKETDAIGQCPYQD 837  
Db 503 TSSRGEBTVRYRDYDHSELTPATTDATG---STROMTWSRYQQLIAFTDCSGYQTRYEYD 559

QY 838 NYNRVIQTIPDGIVNWKVAPFSTDLTIDIRUNGISLG-QQTFDGLSLUTQSQDG-GRV 896  
Db 560 RFGQM-----TAHVRE-----EGISL-YRRVDRGRILTSVDAQGRE 595

QY 897 WAYTWSAGNDQCPSTVTPDGQFIHYQYO----- 925  
Db 596 TRYENAGDL---TAVITPDGNRSBTQYDAMGKAVUSTOGLLTSMEYDAGRVISLTNE 653

Db 1239 VIGWYFLGNGYRVNPVLMRPHPSDMSNPFGRGGINNTYTCGGDPINRIDLNGHLSAGGI 1358  
QY 1330 TGJGMGJAGILLUTAT-GGMATA---AAGTAATAATSTTALAFGLSVTSDFISIV 1443  
Db 1359 LGJVLGAGJIGIVGIVSGLAGAISAGLJAGGAGAINTSALAVTAVIGLAADSGIA 1418

QY 1414 SGALEDASPKASSILGWSMGAGALESAKGGTAKLATHGAFASIDEGENALIKST-SE 1502  
Db 1419 SAIISLESKOPTSGTGLNWIISAGLGVLSFGTSAL-----TFT---SSLVKSARSG 1463

QY 1503 SSRJKWGTSLDRIVEENEEQVIKDHSR 1532  
Db 1464 SQAVSAGVIGSVPLF---GEVASSRSSL 1488

Db 274 AAALTIQPGNGQPAVSKSYEYSSVNFLOGLSGRTSFDSQSONLYLVTKYTSIE--- 330  
QY 357 YDKKEGHQDQTVRIERTINYHILTSECKOONGYIQTETAYAIALIGHFDSQSQFOLP- 415  
Db 331 -RVLDGOSTVSVLERVENKFLMLKBEAKYQDNRKRITTETYNEDLSKSFSEOFENLOOPS 389

QY 416 KTKTETWRSADSNSYRSEITETTEDESIGNFLTKVTKDKKTQKILISPSTHEYYPAGEVDN 475  
Db 390 RVLTRYTDQINTSREBETNIKSFDWGNTL-LTETSGIQK---EVVYVYPNGENS 442

QY 476 CPPRPGFTRFKKIIQTYDSEKFDPKFIOY---SLIGSQ---SIATLKEE 526  
Db 443 CPADPLGFSEFLKSVTQKGSPDAQSVANKVVIHTYQKEPTFTGAYKEYVKVSKSETDN 502

QY 527 RHYSATQUNLNSTLQYNTKSELGRLLKQTECTKGENGKTYSVVHKFTYTKQDTLQSH 586  
Db 503 KIATFSVYNSP---TKSH-GSIAKITSWMNQ---QVTTKVEYSSSEMNTA 551

QY 587 SITTHDNFTIERSQVRSRVTGFLFSDTDTKDIVTOMSYDKLGRULTRLNSCTIYDQYANTL 646  
Db 552 TVTGFDGAMMEKNSKVNTSIVTHRQLRKVDUNHNTDQSYVLIGRITGQIDPGTAREIKN 611

QY 647 YDVEELANNLDDNNRPPFVITPD-----DKWHLMSKSYNWGQJANTHWS 698  
Db 612 VVYQIPGGDENDFWMP-VMLEVDSQGVRKTHYDGMRGRCISIEODDDGAWGWTGSIYQTY 670

QY 699 YTIHQXQYBQGRHHTSYDYLINGRQQT-----DKWHLMSKSYNWGQJANTHWS 753  
Db 671 RKVILARQYDNGQKLSKEISDNWLN---LSANPLVRLATPLVTTKVKYDQGWGNLYSTER 728

QY 754 YGVSEKITVDPITLTAK-----OQSNSNNVQTKGEVTTYTPSQOPIQITLEDEAHL 807  
Db 729 DGRILEIHPVITRITQGVGKGLMUNIQNP-----EQPASKAVYPDGII 776

QY 808 OSCHTILTDRGDWRKETDAIGQCTIYQTYNVRQIQTIDPGIVNWKVAPFSDTILT 867  
Db 777 YSTRTRYDGFGRTVETDAGHATQIGIVDFDRIVKKTLPDGTLIESAYSFHELIS 836

QY 868 DIRVNGISLGQQTPEGLSRITQSODGGRWYATVAGNDQCPSTVTPGQFITYQYPE 927  
Db 837 ALVNGTOLGALVYQGGRITSDTGVGRKTYEYLGQDGL-PIOSITPSHKNKOMDLYY 895

QY 928 LDDAVLQVASNEITQOPSTVNPVFTGALLKAVEGQSLTP- YPSPGRKLMEN-INDMKMS 985  
Db 896 LGSMWSKFTTGTDQDNFRYKISKTGTLUSA-SEGYSQNTSYFSPVQLORESFLDNUKPI 954

QY 986 --YVTLRGLENGYDLTGCTIQTOKISRDTGTRVTOIKDSSKTTIINYDDNHRGSQVTD 1042  
Db 955 SGEYLTYMSGIQLRHKDSFGHNVHSYDQGRLVKEQDAQYATFEYDNGVGLITTKD 1014

QY 1043 LATGMLLTTEFDGSLNREIGGRKLQDSSGHTLDIQQSWLQKQQLANRIVKLYNGVLTQ 1102  
Db 1015 TTSLSQVLTKEIYDADREKRSLSDFSTQV-ITMSYTKNQISORITSDGIVVMKNER 1073

QY 1103 YSDSRNRLNQYKCGAECTDKYGHISIVQNPFTDIYGNITACTHPADGETEDHATK 1162  
Db 1074 YQYDNRQRLSQYQOCGEGOPLDHTGRVQIYHDTQWGNKIRLUNTYRDGKET-VDYD 1132

QY 1163 ANPTDPCQLTTEVHHTPDRMPDNTIRLYKDQAGRGINNT-DHNGNTENFTDITGRL---QN 1218

Db 1133 SQ-ADPTOLIRI---TSDKQIELSYDANG---NLTDRKGQT-LITDQNNRILQVKD 1182

QY 1219 QGGSV---YGDPLNRLVSO---KTDILCCELYRETMVUNVERGEMIRLRTBTIAQ 1273

Db 1183 RLGJLVCYSYQDNLKTAOLANGTUNRO-HYASGKVNOLGIBAATWLNISDQKIGH 1241

QY 1274 ORA---SKVLTGTDQSOSVILTSQDQNLQSEAYGKHKSTANDASIGNGERA 1329

Db 1242 QSAKQGQSVYQYQGHNSTIASONEALMASTYPIYGRSLI---SSLFGLNGAQVP 1298

QY 1330 VSGVTHLGNGRSYDPTMLRHTPDSLSPFAGGINPYSCIGDDINRSDPSGHSLSQWAQ 1389

QY 926 -----PELDDAVLQVASENETTOQPSYNPTVGALLKAVRGQSILPIYYPSGRJKE 976

Db 654 NGSHSHSVFSYDALLRLVQQGPGDRTRHYD-LTGKLTQSEDRGLVILWYBDSRTHR 712

QY 977 NINDMKKNSYLWTGLTGLENGTYDGTGKISRDTGHVTOIKDS---IKITLNTDDL 1032

Db 713 TNGBPAEOMQY-----DGHGMILTDISHLSBCHRVAHVGYDDK 751

QY 1033 NRHIG--SQTDLAIGHM-----DGHGMILTDISHLSBCHRVAHVGYDDK 1049

Db 752 GRJTGECOTVNPETGELIWIHQHSTKHAYNEOGLANRVTPDSLPPVEMLYGSGYLAGMKL 811

QY 1050 -TTVERP--DGLARREIGRKLCDCSSH--TLQIQQSWLKIQOLANR----- 1089

Db 812 GCPPLVEVTRDRLHRETVESFGMSNAGSNAYEJSTSYTPAGQLOQSHNLIVYDRDGWS 871

QY 1090 ---IVKUNGVLORTEQSYDSRNRNLNOYKCDGABC----PTDKYHSI----- 1130

Db 872 DNGSLVLRISGPRO-TREYISATCLESVRTLAPDLDIPIPVDPAQRNLDPDEHLHDS 930

QY 1131 -----VTQN---FITYDIGNITACHTFADG---TEHDATPKCANPDPQCQTEV 1174

Db 931 TLTWMPNDRIADEAHYVTRHBYGRLTEKDTRIPAGVIRTDERT----- 975

QY 1175 HHTHPDMPDNIRLUKDKAGRVINID-NHGN---TENITYDINGR----- 1215

Db 976 HPH-----YDSOHFLVFTRIQHGPPLVEVLYDPLIGRMKVRWERRDITG 1025

QY 1216 -----LQNGOGSISVGYDPLAHLVQSOQRTDUDCELYR----- 1247

Db 1026 WMSLUSRKRKVUTWQWDG-DRITVQTDTRIQTWEPGSPTPLRVETENGREKAQRS 1084

QY 1248 -ETMLVNEVRG-----EMRLL-RFGETIA----- 1279

Db 1085 LAETLQOEGSENHGHHGVVPBALKVLLDRLEEBRADRSVSESSRAWLAQCGLTVEQALARQV 1144

QY 1280 LILTQDSDQS-----VILTSKONLSQEA-YSAYGKHKSTANDASILGYN--- 1323

Db 1145 BPEYTPARKAHYTHCDHRGLPLSBDGNTAMSAYEWBGNGQNLBEPNPHV--YQPTRL 1202

QY 1324 -GERADPVSGVTGTHLGNGYRSYDPTLMFRTPDSLSPFG-AGGTNPVSYCLGDPRINRSDP 1381

Db 1203 PGOHONDESSLYY--NRHYYDPLQGRYITQD--PMGLKGHNLYQPL-WFLQQLDFM 1256

QY 1382 GHISWQAMTGGMGIAGLTLTATGGMATAAAGGIAAAIASSTTALPAGA-SVTSIDS 1441

Db 1257 GLL---QTWDDARS-----ACTYQGV---CCVLSLRTGPSKDSTADAALD----- 1296

QY 1442 INSGALEDASPKAISLIGWVSMGMAAGLAESA1KGKLTATHGAFADGEJNLLKSTS 1501

Db 1297 ---ALKETONR-----LCDNEYSGIVCKOTG-----KYFASKAEDNLR--K 1336

QY 1502 ESRRIKGWVTRSLDREIVENEGOVIKODHSRGYTDNFMGKGEQALILVHGDK--DGFLYHT 1559

Db 1337 E\$VPLRKRCPTGCDRVAHYTHG--ADSHGDYVDEFSSSSDNKLVLVRSKDNNEAFYLT 1393

QY 1560 EGNKH--NGKGPY 1570

Db 1394 PDGRPEALLNKGEY 1407

RESULT 3  
US-09-711-164-467

; Sequence 467, Application US/09711164

; Patent No. 6589738

; GENERAL INFORMATION:

; APPLICANT: Forsyth, R. Allyn

; APPLICANT: Ohlsen, Karl

; APPLICANT: Zyskind, Judith

; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTI-SENSE THERAPY

; FILE REFERENCE: ELTRA-008A

; CURRENT APPLICATION NUMBER: US/09/711-164

; CURRENT FILING DATE: 2000-11-09

PRIOR APPLICATION NUMBER: US 60/164415

PRIORITY FILING DATE: 1999-11-9

NUMBER OF SEQ ID NOS: 469

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO: 467

LENGTH: 1377

TYPE: PRT

ORGANISM: Escherichia coli

US-09-711-164-467

Query Match 3.7%; Score 331.5; DB 2; Length 1377;

Best Local Similarity 20.8%; Pred. No. 3; Se-16; Gaps 75; Matches 288; Conservative 170; Mismatches 450; Indels 477; Gap 75;

QY 211 LPGQKCEY-RTSLPRLRNLQNSIHNFSLGNENPLTISFGYPIKGKQGILGQWITSMAP 269

Db 126 LFPGBR-DCYVSRSLSLWLR--SGVAKUBGHRALALWQ----- 160

QY 270 GGKJETWYS-----NNNQG-----HRPQSAML--PVLPTVILMKQVPGACQ- 310

Db 151 -ALPEELRLSPHRYLATNSPQGWMWLGMCERVEPEALVPLPPVTLGIVDRCRT 219

QY 311 PAIQAEVSYTSHNYVGEGCSNGINNNKUNLNLGYMLGTYTNGSTERSRYDKEGHDQIVIE 370

Db 220 QTPHREAMGFBFSCHITGIVDAMRH---FRLVL-----TQAQR----- 255

QY 371 RTTNYHHLITSEKQONGYIQTETAVVAILGNHFDSCPSQPOLPKU--KETWRSND 427

Db 256 -----AEARQO-----AISG---GTPSAF--PDTLPGYTYEGR--DN 287

QY 428 STSBSTTTFDESGNPITKVKDKKQKISPTSTHWYYPAGEVDNCPPB--YQFT 484

Db 248 GIRLSAWMLTHDBP-----Y-----ENLPAAPLVRYWT 317

QY 485 -----RFVKKIQTOPYSBEFK-----DDEKFQIORYSLSLGSOSHVTKIE 525

Db 318 PRGELAVYDRSQKQVRSPTYDDKQYGRMVWAHRTGPE--JRYD--SDGRVTEQLN 372

QY 526 ERIVSATOLNSTLFQNTDKSBLGKRIKOTSC--TKENGKTYSYVHFKFTVKQDDPLQ 583

Db 373 PAGLSYLT-----YQEKDRITITSDLRREVLTQOCBAG-LKRVWTK----- 413

QY 584 QSHSITHDNFTHSRSQRTYGRFLFSDDTKDIVTOMSYDKLGRJLRTLNGTYAN 643

Db 414 -----ERADGSYUQSQDPA--VGLRQMTDAGRTR-BYSPVYTGJLITR--TPQGR 462

QY 644 TLTYDYEMLQDNRRPPFVITTDVUNQNLNEFDGAGRHSQCLKSD---GDKP 698

Db 463 ASAFYINHNO-----LTSATGPDQBLRREYDELGRJLQTAQDGDITRYDPMH 514

QY 659 YTHTQOQDDEQPHRTSTSVDLNGRQTDPKVHMSMSKSYDNWQGQIANTHWSGYSE 758

Db 515 SDLPCATEDATESRKTWMSRY--GQLSFTDCSGVYTRYDDEDRFGMTAVAREGILSQ 571

QY 759 KITVDPI-TITATKQLOSSNNNVQTKVTTPSQQPIQTLFDEAQHQ-----SC 810

Db 572 YRAYDSRQOLIAVDTQCHETR-----YNTAGDLTAVAPCDSR 612

QY 811 HTLJRDGDRVRKETDAIQCITYQDNYNRVQIQTLDGTTVNRKJAPFSTLTLIDR 870

Db 613 NGQYDANGKAVFTQG-GLTRSMYDAGRVRILTSNGS-----HTPFRVLDRLHQ 666

QY 871 VNGISLSGQPF--DGLSLRUTQSGDGRVWAYTYSAGNDQCRSTVTPGQFTHYQYOPBL 928

Db 667 ETGFDGRTRQHDLTGKLRSEDEGLV-----TRWHD-BA 702

QY 929 DDAVLQVASENETTOQPSYNPTVGALLKAVRGQSILPIYYPSGRJKE 981

Db 703 DRTHRTWKGTERWQD-BRWLTDISHLSBCHRVA-HYTYDEKGRRLTBQRTVHP 760

QY 932 KRNQSYL-----TLEGLENQTYDGTGKISRQHTRGRTQIKOSSIKTLYDILN 1033

Db 751 QTEALLWQHETRAVAGLAN--RCIPDSLPAVEMLYG-----SGYLAGMKGUD-- 809

Qy 1034 RHICSOVTDIATSHMLTTTVEF--DGLREIREGRKLCDSGGHTLIDQOSWLKOOLANRIV 1091  
Db 810 -----TFLVEYTRDRILARETLR----SPGRVELTIAVTPACOLOSO-- 846  
Qy 1092 KNGNLQRTEQYSVNSRNLNQYKDG&CPTDKYHSIVTDFYDYGNIATCHTEA 1151  
Db 847 HLNLSLS-DRDYTWNDNGBLIR----TSSPRO-----TRSYSYSTGRLAVGHTAA 893  
Qy 1152 DGTEDHATFKANPTDPC--OLTEVHTHPD---MPDN----IRLKDKAGRVIN 1197  
Db 894 ----NLDRIPVATDPAERNLPL-PELRPDSTLSMWPDRNARDAHYLYRKHGRATE 947  
Qy 1240 LDCEL-----YY-----RETMVNE-----VRNGEMILR---TGTI 1270  
Qy 1198 ITD-----NIGHTENTPYTGLGRONGQ-----GSVYGYDPLNRLVSQKDT 1239  
Db 948 KTDLPIPEGVIRTDDEBTHRHYDQSQRHLHYTRIQEEPLVESKYLIPGRVAKWR 1007  
Qy 1271 IAQRA-----SKVLTGTSQSIVL-----TS 1294  
Db 1068 KTOEISLADALQSGGEDGSVWFPVLUQMDLESETLADRVESEBSRRWLASCGLTVE 1127  
Qy 1295 DKON-----LSORA-----VSAYGRHKSTAND--AS 1318  
Db 1128 QMQNQMDPVYTTPARKIHLHYDHRGLPLALISKGTETCAEBWGNLNEENPHQIQQ 1187  
Qy 1319 ILGTYGERADPPVSGVTHLNGYRSYDPTLMFRHDPSLSPPG-AGGINPYSYCLGDPINR 1377  
Db 1188 LIRLPGQQYDEBESGLYY--NRHRYTDPLQGRYITQD--PIGLKGWNFYQYPL-NPVTN 1241  
Qy 1378 SDPSG 1382  
Db 1242 TDPLG 1246

---

RESULT 4  
US-09-252-991A-28424  
; Sequence 20424, Application US/09252991A  
; Patent No. 6551195

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenstein et al.  
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252, 991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074, 788  
PRIOR FILING DATE: 1998-02-18  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33,424  
SEQ ID NO 28424  
LENGTH: 998  
TYPE: PRT  
ORGANISM: *Pseudomonas aeruginosa*  
US-09-252-991A-28424

Query Match 3.6%; Score 317; DB 2; Length 998;  
Best Local Similarity 22.0%; Pct. 2.7e-15; Matches 222; Conservative 129; Mismatches 362; Indels 296; Gaps 47;

Qy 504 EKPKIQYRYSLSQSNTVKKEERHYSATQDNLSTFQMTDKSPLGRLLKQTCIGEN 563  
Db 112 ERAWSYSYNAG-----LTERADCPRTDWDVTLAYD-----SRGNL 149  
Qy 564 GKTYSVNHKFPTYKQDQLQOSHSTITHDNTIHSQVRSRYTG--RLFSDDTKDIVQ 621  
Db 150 TQVNLALGQVTRLGQYDVERGKPGSTDANGTSSLA----YTGVDMGLASVAGSTR 204  
; SEQ ID NO 28424  
; LENGTH: 998  
; TYPE: PRT  
; ORGANISM: *Pseudomonas aeruginosa*  
US-09-252-991A-28424

---

RESULT 5  
US-09-328-352-6604  
; Sequence 6604, Application US/09328352  
; Patent No. 6562358

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTCS9-03PA  
CURRENT APPLICATION NUMBER: US/09/328, 352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252

Db 205 FDYDAVGQTTRVTRGDG---SWLSYEX-----DDAR--LVAIGNNLGERLYDVDTK 252  
Qy 682 GRVSOCKKDSQDGKTYTIFHQ-Q-YEQR-----HETSTY-S-DYLNGRQDTP 730  
Db 253 GNETAQKRDASG-----LVRQOWATBBLGILURAVGAGGOTRSFVDLNPNVGP 308  
Qy 731 DKVHLSMSKSYDNWGQIANTHNSYGVSEKITUDPILATKOLQOSNSNNVQTKEVFTT 790  
Db 309 RQ--FANSQAFDALDRUGQSPLGAKTRVLAD-----AQDNITEVKOPRGVTY 357  
Qy 791 PSQOPTQITLFDEAGHL-----QSCHTLTRDGWDRVRKETDAIGOCTIYQDNYMVTI 843  
Db 409 ERNSP-----SDPSLIVQYRDLTADGNOGIGRLGIAEGARDS----- 446  
Qy 904 GNDQCPSTWITPDQOFIHYQKPELDLAVLQVASNETTOFSYNPVTCALLKAVAEGL 963  
Db 447 -----LTYRVO-ERGMNVEQPSRISRLDQTLDRVTRYDA---NQL 485  
Qy 954 TPIYPSCLRKMENINDKMSYLWTLURGLENYTDLTGTOIKISRDRTHGRVTOIK--- 1019  
Db 486 LEIGPS-----GLATCY-----PRNGQQVASVTAVG 514  
Qy 1020 DSSIKTT-----LNYDDLNRHGSQVTDLANCHMLTTVEFDG---LNREIGKUDS 1069  
Db 515 DKAPSTLVQIAVLPFGSLQR-----LTWNGITISREVNQDYQILRQV----- 560  
Qy 1070 SGHTLIDQSQWLKQQLANRIVKLNGVQ-----TEQYSYDSRNQYK-CDGAE 1121  
Db 561 -----PWQSDYQ----HDANGNIOQHRSHLWMGTFDQYDPLDPLRULBERGVQGC-- 604  
Qy 1122 PTDKYGHSLVTONFTYDIFTYGNIT-ACHTFADGETDHATFKEAMPTDPCQLTETVHTHPD 1180  
Db 605 -----RESYAYDAVGNTVQRSDNPASGCTTASSDQYQ----- 636  
Qy 1181 MPDNIRL-----KYKAGRVTINTTNGHTENFTDTGLQN---GOCSVXSD 1227  
Db 637 -PDSNRLTAIGAQATVSDAAG---NLTDODRA-ARKLAIDAQGLQSVSLDQOVAEYIN 691  
Qy 1228 PL-NRLVQOKTDLDCELYRETMLVNEVNGEMTRLTG-----ETIAQRA 1276  
Db 692 ALGGRIVKLTPESTITYLPGDQGLGKAEDGSGKRQYQYIWLDSLPLATDADYD 751  
Qy 1277 S-----KVLTGTSQSIVLSDK--QNLQBAYSAIYKHKSTANDAS--LIGINGER 1326  
Db 752 QGKVGNPNTLWYHGDLHPTPLATDASQIAWQWQSDAFGRGRALSQSTQVNLFPCQY 811  
Qy 1327 ADPVSQVTLNGYRSYDPTLMFRHDPSLSPSLSPRG-AGGINPYSYCLGDPINRSDPSG-- 1382  
Db 812 YDAESGLY--NFTRDYDPETGRYEST--PIGLSGSYNTGYVQGAPLNRLDPLGAA 866  
Qy 1383 --HLSWQAWTGCMGI--AGLUTINTGGMATAAGGIAAIASTST 1425  
Db 867 IEIDIPKSAWDWIPGNRUPAGRILL---GGVLLIVASISGATFOADSIT 911

LENGTH: 1627  
 TYPE: PRT  
 ORGANISM: *Acinetobacter baumannii*  
 US-09-328-352-6604

Query Match 3.5%; Score 313; DB 2; Length 1627;  
 Best Local Similarity 20.4%; Pred. No. 1,3e-14; Mismatches 506; Indels 498; Gaps 73;  
 Matches 303; Conservative 177; Mismatches 506; Indels 498; Gaps 73;

Qy 75 IGRGIGFNGLGSVDRKRSNLLSLSTGENY--KVIEKDVKV-----KLO 115  
 Db 369 ITFMAGTTRPFTHDAQLQGILQDLSRISRTVNSVNUQMDAIFGARMWPTTTSRKPCYT 428  
 Qy 116 QKQDNLPRPEKDJKENCYRITHKSGDIEVLTGFFNNNAFDLKVFKKLNPAGHAY--ID 172  
 Db 429 SKKKDKHKOYLNGLYVIC-----LJGRALDLPDIJK-----GSIVDPBQ 468  
 Qy 173 WNFTEATOPRLNRT-----YDQDGHDIPILANLEYGLKILITLPGQEGYRTEL 224  
 Db 469 YTYTVLSPDOLHLIAYGEDBKRYYKGYDYLSTIE-----RNKGPKVALR 514  
 Qy 225 R----LNRQLNISHNFSLEGNNPLTWGSGYPTG-----KNGLQWMTSMTARG 271  
 Db 515 YDNVSTDNKTCILSDLPKDDNNLAHLAQTLQGILQDLSRISRTVNSVNUQMDAIFGARMWPTTTSRKPCYT 568  
 Qy 272 LKETVUNSYNNQGHIFPOSANLPVLPYVTLMKVPGAGCOPAOEBSYTH-----NYV 325  
 Db 569 -----YDIDQOG-----DLVQATNEPASYYVYQTHHLITYTRYDYL 604  
 Qy 326 GGSSNGINWNKLNLQYGLMTEYNGSTESSRRVQKEGHDQIVIERTONYNLTSEKQ 385  
 Db 605 HGRKMLKNDGILPPTSKAIBEWADNASRASKLEMKD-----NIRKT-----TVDLVEGNS 653  
 Qy 386 QNSYIQTETAYAALIGNFPDSQSQFOLPKTKTETWSADNSYRSEITETTDFESGNPL 445  
 Db 654 TERYIDYDGYTYRIVPNE---ECFPRDDAKNITHIAKOSK---TSYTYDEGRNL 707  
 Qy 446 TKVJDKKKTQKIIISPSTWAEYYP-----AGEVDNCPPEPY---GFTRPVKKCIQTPVNSE 498  
 Db 708 TTQDGGAT-----SFYEDEGNQLTOMDABQRGRWKQDQSGNLKEI----- 752  
 Qy 499 FKDOPBEKFQYRSLIGQSOSHVTIKERRHYSATQOLLNSTLFOYNTDS---ELGRLIK 554  
 Db 753 .-----DPLK-----HETAVV-----YNAAGLVTISITDAKGGSKSLKVDQGNLIS 792  
 Qy 555 QTSC-----TK----GENGKTYSVH-----KFTYKQDDTILQO----- 584  
 Db 793 YTCGSKETKWDYERGRVISTENALNOKVEYFETLLENREPIKGLPLNAGQLEKI 852  
 Qy 585 SHSTTHDNFTIHSQVSRTYRLFSQDTDTKDVTONMSYDUGRLRTRT--LNSGTPYA 642  
 Db 853 KHDGTEEFH-INDAE---GRLLAHVDPKONITRYEDBAGLILSRDALM----- 899  
 Qy 643 NTLYTDYB-----LNLQDDNRPRPVITDVGNSQLRMEPDGAGRHYSOCLKDSGDKF 698  
 Db 900 HKUKYKWDRLGRTRLINE-----NGASYQFYDVASRIV---KEBDFDKE 943  
 Qy 699 YTHHTQOQDEQ-GRHHTS---TYSQYLNGRQQDTP-DKFLHMSKSYNQGQIANTHWSY 754  
 Db 944 TVTH---YDEKSQOLATLEVASYAGQDLPKDRAPKDMIQQT-----DSMGHLFQRTAGY 997  
 Qy 755 G-----VSEKITVPI-TLTATKQLOQSNNSNVQTCKEVTTYTSQQPIQITLDEA 804  
 Db 998 GHYGLELBKQKTBEPFAYDMGMRITOQAKQAQNSIQ-----WFDAA 1037  
 Qy 805 GHQOSCHTLLTRGDMDVRKETDAIGQCTI--YQDNYMRVQITLPDCITVNRKYAPEST 862  
 Db 1038 GNLVOEH-----QDQDINKINTAWMKHQDBINDRIKTRPDSQVID----- 1078  
 Qy 863 DLTUDDIRINGISLGQOPTDGLSLRUTQSGDRGMWATYSAQHOCQSTVTPDQFQHY 922  
 Db 1079 -----W-LTYSGSHVQ---SLINGQDFVSP 1100

Qy 923 QYQPELDDAVLQVASNETIQQSYNPVTFGALLKAVAEQGSQSLPIYVPSGRKLMENINMK 982  
 Db 1101 ---ERDDLHRETA-----RHYANGVSQEQQYDLAGRLKSQM----- 1134

Qy 983 KMSYLWTWTRGLENGYTD-----LNTQIOKSR-----DTGVRVQIKDSSI-KTINY 1029  
 Db 1135 -----1SEHENHGQYQKRNNALEQTSQVQLVQYDQKTYGKBLTARDTRRGNIAKY 1187  
 Qy 1030 DDANRHGSQVQDLATGMILTTVERDQGLNRETRGKLUCCSSHTLDQOSWIKTQQLA----- 1087  
 Db 1188 DPVGRLLB-----ASSLGKETPSFDPASNL-----DSYHSOKYQSHSGQKUDERSYG 1235  
 Qy 1088 -NRIVKUQGVLSR-----TBYQSYDSRNRLAQYKCDGAECPDKYHSIVQNFTDIYGINIT 1144  
 Qy 1236 YNLVU-NNVVKYLDQGQXQDYAGQLRQKTSQGDL-----NLEWDYQGRMV 1281  
 Qy 1145 ACHTFAGCTEDHATPKFANPPTDPCQJLVEHHTHPMDPNIRLKDYKDGKRVNITDNGN 1204  
 Db 1282 KSRNS-----QYTAEY-----RYDALGRR1QKWSKHH 1309  
 Qy 1285 TENFTYDILGRQNGQSSVYGDPLNKVUSQKDTLOCBLVYRETNUENEVNGEM--IR 1262  
 Db 1338 LKQTPDMSDRPMIHRPLWKTKEKEGFEDWMFHQDHGLPQEMWHTGAIWKAEK 1417  
 Qy 1310 T-----GQBN-----1IYGWDG-DTIAVYESTEELTHRYIYEDKFSPVPMLOAVVYSPIE 1357  
 Qy 1263 LLRTGETT-----IAQORASKVULGTDSQSQSVLTSKDKNLSQB-----AYS 1305  
 Qy 1306 AYRKHKSPANDASI-----LGTINGERADPVSGCVTHLNGYGRSYDPTLMRFHEDDS 1355  
 Db 1418 AWBCKAKBAAKSYPFENSESIIISNIRFGQOYFEEBTGHY--NRYRISPYVGRFVSKD 1474  
 Qy 1336 LSPFG-AGINPNSYCDPINSRSPSCHLSQPAWTGIGMGMGAG 1398  
 Db 1475 --PIGLLQGNNVYVA-KNPITWIDSKGKLCSTTLNRNIG-GVKG 1514

RESULT 6  
 US-09-543-681A-5329  
 Sequence 5329, Application US/09543681A  
 Patent No. 6605709

GENERAL INFORMATION:

APPLICANT: GARY BRSTON  
 TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
 FILE REFERENCE: 2709-1002-001  
 CURRENT APPLICATION NUMBER: US/09/543, 681A  
 CURRENT FILING DATE: 2000-04-05  
 PRIOR APPLICATION NUMBER: US 60/128,706  
 PRIOR FILING DATE: 1999-04-09  
 SEQ ID NO 5329  
 LENGTH: 1586

TYPE: PRT  
 ORGANISM: *Proteus mirabilis*  
 FEATURE:  
 NAME/KEY: UNSURE  
 LOCATION: (1576)  
 OTHER INFORMATION: Identity of amino acid at the above locations are unknown.  
 US-09-543-681A-5329

Query Match 3.4%; Score 306; DB 2; Length 1586;  
 Best Local Similarity 22.1%; Pred. No. 4,4e-4; Mismatches 385; Indels 342; Gaps 59;  
 Matches 243; Conservative 129; Mismatches 385; Indels 342; Gaps 59;

Qy 499 FKDOPBEKF-----IQYNS-----LIGSQSHVTIKERRHYSATQOLLNS- 537  
 Db 486 PEANPKINTLRLMSADRNHNAHYRKAENDGELVQTHDAYDTRDLYHDTORLQSV 545  
 Qy 538 -----TLCQYNTKSESLGRLLKQCTCTKGENGKTYSVHKTFTQDFTL--QQS 585  
 Db 546 TRHQGQEKTLLVYTYDQ-Q-RIVQVNAKD-----RVTRFEGMDDESSIMAMHQA 596

GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 7560  
LENGTH: 1439  
TYPE: PRT  
ORGANISM: *Proteus mirabilis*  
US-09-543-681A-7560

Db 970 DLHREIS-----RTOCBLTQYRQDCKGRTIYSTPSRDQKHLINGITILWKRKFDP----- 1021  
 Qy :|: ;|: ;|:  
 Qy 1088 NRIKVKGVLVLTQTEQSYDSRMRLN----- 1022 GNLGAMEDTYRGWVLYDSEORLKKVASSENIDAMLFYDRADNLLERPOSEMDAEHSPT 1081  
 Db 1124 -----DKYHSIVTONFTKIVGNITACHTTPADSTEDHATFKRANPTDFCOLTEHH 1176  
 Qy |:|: ;|: ;|:  
 Db 1082 LELSPCGDKL-RQFOGMHYQDAYGNIA----- 1177 THPDMPDNIRLKXDKAGRGRVINITDNGNTENITYDTGR 1176  
 Qy |:|: ;|: ;|:  
 Db 1117 1177 THPDMPDNIRLKXDKAGRGRVINITDNGNTENITYDTGR----- 1215  
 Qy |:|: ;|: ;|:  
 Db 1121 -----AYGDGRLVIAHNGIKAQHRYDALGRKHTVNRKESQVRQETWF 1168  
 Qy |:|: ;|: ;|:  
 Db 1216 -----LONGSGSVY----- 1216 1216 -----LONGSGSVY-----GYDPRNTRLVSQKTDICBLJYRETMLVN--- 1253  
 Qy |:|: ;|: ;|:  
 Db 1169 IWQGLRLLQBODINTGKHTOCYCEBGSYTPLAVIKQSSGF--- 1304 1169 IWQGLRLLQBODINTGKHTOCYCEBGSYTPLAVIKQSSGF---RYRN-----QSSQY- 1120  
 Qy |:|: ;|: ;|:  
 Db 1254 EVRNGEMIRLRTGETIAQO----- 1254 RASKLIGTDSQSOSVILSDKQNSQRAYSA 1306  
 Qy |:|: ;|: ;|:  
 Db 1225 BVTNAG-----GNTIWSKGYERFGRVRSPLSPVSPER-- 1225 1225 BVTNAG-----GNTIWSKGYERFGRVRSPLSPVSPER--VMASSEQUR----- 1253  
 Qy |:|: ;|: ;|:  
 Db 1307 YGKHKSTANDASILGNGERADPVSQVTHLGANGYRSYDPTLMRFPHTPSLSPFG-AGSIN 1365  
 Qy |:|: ;|: ;|:  
 Db 1269 -----YAGQFDNEIGL-HP-NTRRFDPOIQKRPIMP----- 1308 1269 -----YAGQFDNEIGL-HP-NTRRFDPOIQKRPIMP-----PIGLGGIN 1308  
 Qy |:|: ;|: ;|:  
 Db 1366 PYSYCLGDPINRSDPSG----- 1366 HLSNQAWTCI 1392  
 Qy |:|: ;|: ;|:  
 Db 1309 LYQVA-PNPLCWIDPNGLSSOBMVRVRHTSVEGGLEGI 1345  
 Qy |:|: ;|: ;|:  
 Result 8  
 US-09-252-991A-23805  
 ; Sequence 23805, Application US/09252991A  
 ; Patent No. 6551735  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSUEDOMONAS  
 ; FILE REFERENCE: 10716-136  
 ; CURRENT APPLICATION NUMBER: US/09/252.991A  
 ; PRIOR APPLICATION NUMBER: US 60/074, 788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094, 190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 23805  
 ; LENGTH: 1626  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-23805  
 Query Match 3.3%; Score 292; DB 2; Length 1626;  
 Best Local Similarity 21.1%; Pred. No. 5.7e-13; Matches 204; Conservative 125; Mismatches 316; Indels 320; Gaps 46;  
 Qy 597 HRSQVRSRY---TERLFSDDTKDVTOMSYDKLGRLLTRTNTSGPY----- 641  
 Db 633 HTTQLARYKCNERGQLIBAGNALLVARRYDARAVILEROLAGGASFFWWEGECKQA 692  
 Qy 642 -----ANTLYTD--YELNN----- 642 642 -----ANTLYTD--YELNN-----LQDDNRPFPVTTDVNGQLRN 677  
 Db 693 RAVHWHASFPOMDSRKYWNEDGSGVTAINDGSEEVYVHDNAR--LVRQVDPDGGETLRH 750  
 Qy 678 FDGAGRRHVSQCLKDGGDKGYTIHQYQDQR----- 718 678 FDGAGRRHVSQCLKDGGDKGYTIHQYQDQR-----HHTSTY----- 718  
 Qy 751 YDEKOLVVAS-RDPLG----- 751 YDEKOLVVAS-RDPLG-----TITYRYDQAGRALLPAGEPCTCYSDFGFVSRV 803  
 Db 719 -----DYLNTGR---QQTDD--KWHLSMSRSYDNGQIANTHWSVGSEKITVDPITL 767  
 Qy 804 GEAQMYKVERNAQGD-TROTDPFEGNVTHY---AYDHCGLVETIHADGSLHOLTMFLGQ 859  
 Db 860 LIEQDGDGS----- 860 LIEQDGDGS-----VRYRVPYDGTQQTRODSEGAI----- 891  
 Qy 828 IGQCIYQYNTNRVIOITHPDGTIVNKWPASTFDTLTDIURVNGISLGOOTPDFGSLR 887  
 Db 892 ---TRFWMDAGRSQUTIPGJ-----ASRMY-----RYN-----AYGKV----- 924  
 Qy 1082 1082 1082 LELSPCGDKL-RQFOGMHYQDAYGNIA-----RYRN-----QSSQY- 1120  
 Qy 1117 1117 1117 THPDMPDNIRLKXDKAGRGRVINITDNGNTENITYDTGR-----RYRN-----QSSQY- 1120  
 Db 1121 1121 1121 -----AYGDGRLVIAHNGIKAQHRYDALGRKHTVNRKESQVRQETWF 1168  
 Qy 1216 1216 1216 -----LONGSGSVY-----GYDPRNTRLVSQKTDICBLJYRETMLVN--- 1253  
 Qy 1225 1225 1225 BVTNAG-----GNTIWSKGYERFGRVRSPLSPVSPER-- 1225 1225 BVTNAG-----GNTIWSKGYERFGRVRSPLSPVSPER--VMASSEQUR----- 1253  
 Qy 1307 1307 1307 YGKHKSTANDASILGNGERADPVSQVTHLGANGYRSYDPTLMRFPHTPSLSPFG-AGSIN 1365  
 Qy 1269 1269 1269 -----YAGQFDNEIGL-HP-NTRRFDPOIQKRPIMP----- 1308 1269 -----YAGQFDNEIGL-HP-NTRRFDPOIQKRPIMP-----PIGLGGIN 1308  
 Qy 1366 1366 1366 PYSYCLGDPINRSDPSG----- 1366 HLSNQAWTCI 1392  
 Qy 1309 1309 1309 LYQVA-PNPLCWIDPNGLSSOBMVRVRHTSVEGGLEGI 1345  
 Result 9  
 US-09-328-352-5542  
 ; Sequence 5542, Application US/09328352  
 ; Patent No. 6562958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GUY L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; FILE REFERENCE: GTC-99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328, 352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 5542  
 ; LENGTH: 1596  
 ; TYPE: PRT  
 ; ORGANISM: Acinetobacter baumannii  
 ; US-09-328-352-5542  
 Query Match 3.2%; Score 285.5; DB 2; Length 1596;

Db 1194 GIVGVGRRLYKTKAELTLEFGMDGGDDMIIWESFKSAQNTYKHYIYBDSFVFLQLQAGYKD  
Qy 1343 YDPTLMLRHTPDSLSPPGAGGINPYSSYCLGDP1-NRS 1378  
Db 1254 P---IQLETPD--YQEYQTKPIS-IYKDPFWNRRN 1282

RESULT 10

US-09-252-991A-23640

; Sequence 23640; Application US/09252991A

; Patent No. 6351795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAL

; FILE REFERENCE: 10196 136

; CURRENT APPLICATION NUMBER: US/09-125, 991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074, 788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094, 190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 23640

; LENGTH: 974

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23640

QY 1084 QQLANRIV-----KLNGLVOR-----TEQYSNDSRNRLNQYK-CDGACPCPTDKY 1126  
 Db 552 YQLLQRQKCPWQSYQHDAANGNQOHRHSIWMGTDYQDPLDLTTEERGVGG----- 604  
 QY 1127 GHSIVTONTYDYGNT-ACHTFADGPEDHATPKFAMPTDPQCQLETVHHTHDMPNTI 1185  
 Db 605 -----RSYADAVGNRORSDNPASGGTASSQDYOA-----PDSN 640  
 QY 1186 RL-----KYDAGRGRVNITDNGHNTENFTVTLGRQN-----GQGSVYGDPU-NR 1231  
 Db 641 RLTAAQGAQAVTSDAAG--NLTOPIA-ARRKLAQDAQGRILQSVSLDGQOVAEYRNALGQR 696  
 QY 1232 LVSOKTDTDCELYRETMVNEVRNGEMTRLLTG-----ETIAQOBAS---- 1277  
 Db 697 IVKLTPESVTYLYGPGDQIGEAEHDSGRKLRQAQYVWLSPLATIDAYDAQGRV 756  
 QY 1278 -KVLTGTDQSOSVILTSKD--ONLSQRAYSAVGKHESTANDAS--LGYNCRERADPV 1331  
 Db 757 NPTILYLGHDHDTPLATDASGQIAWQOMQSDARGRGEALSQSIQVNMRPGQGYDDE 816  
 QY 1332 GVTILGNGYGRSYDP-----TLMRPTHD-----SLSPPGAGGIN-- 1366  
 Db 817 GLRY--NYFRDVDPBTGRYVESPDEVTLRKLNPNEMTFLNSGSMLOAQTPWHSHGFTNH 874  
 QY 1367 -YSVCLGDPINRSD-----PSGHLSW-----QAWTGIGMIGAGLTTAT 1405  
 Db 875 NYTTSNDNNPTAKSDPKHGLSPNPTNLYVPTDNTCTLECKKRFITGKSF----- 925  
 QY 1406 GHMATAAAGGIAAMIASTSTTALAFGALSUTSDTISIVSGALDASPKA 1456  
 Db 926 -----LvgalCskRATTFF-FGGVVCNSTIVMCCASCSOCNRAPS 965

RESULT 11  
 US-09-543-681A-4476 Application US/09543681A  
 ; Sequence 4476 Application US/09543681A  
 ; Patent No. 660709  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GARY BRETON  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/543,681A  
 ; CURRENT FILING DATE: 2000-04-05  
 ; PRIOR APPLICATION NUMBER: US 60/128,706  
 ; PRIOR FILING DATE: 1999-04-09  
 ; NUMBER OF SEQ ID NOS: 8344  
 ; SEQ ID NO 4476  
 ; LENGTH: 1665  
 ; TYPE: PRT  
 ; ORGANISM: Proteus mirabilis  
 ; US-09-543-681A-4476

Query Match 3.1% Score 271.5; DB 2; Length 1665;  
 Best Local Similarity 20.7%; Pred. No. 2.4e-11;  
 Matches 233; Conservative 148; Mismatches 350; Indels 395; Gaps 66;

Qy 508 QYRYSLGQSOSHVLKIBERRHYSATQLNLSLQYNTDKSELGRALKQT-----EC 558  
 Db 608 QHRDIPITNEQSOQSV-KRHDGALIL---BYQOK--LURICSISSTPSPDFVRC 659

Qy 559 TKGENGKTYSVVHKFTYTKODDTIQOSHISITTHDNFTIHSQRSRYGR---LFSTD 614  
 Db 660 EYDTHG-----YLSQCHAVQNH-----LWHRYPEGIMAWGTD 695

Qy 615 TKOTVTONSYDQKGRLLTETLNSG-----TRYANTLY-----DTELNIQDDNRP 661  
 Db 696 STELT-IDYBEGQRVATHPSGFWNDRFIDYQRMFTYDAGEGFSRYVYDDN-- 750

Qy 662 FVITT-----DVMGQLNREFGAGRHSQCLD--SPDGKFY 699

Db 751 -LVRTIDBLWRETYTEMQKKLAEINIGERBY--GYHNGLGLVYIYLPGKQIY 806

QY 700 TIHQQYMBQGR--HHTSTSVDYLNGRQQTDPDKVHLSMSKSYDNMQOIANTHWSYCWS 757  
 Db 807 -----DINDYQGLTHFTSAGD-----EWOLSYDENQNL----- 836  
 QY 758 EKITWDPTPLTATKOLQNSNNYVTGKWT-----YTSQOPIQITLPDEAGHQS 809  
 Db 837 --IVTDP--OGROQVYBSOHGBLKAITPNAQWQHYNPAHOLIKT--NPYQHSTE 889  
 QY 810 CHTIRDQMDVRKGTDAGQCTYO-----YDNYNRVI 843  
 Db 890 YHS---DEIGRLAHYTDALHTRYOKSTEHASTNGSISKILLPDGVHQOIDTSEERVI 946  
 QY 844 QITLPDGIVNPKVAPPSTDMLTDIIRNGISLGQQTGDSRUTQSD-GGRWAWTY- 901  
 Db 947 AVTDGEGKTRKRYGPF-DVLLAMIRDGSEI-RFDPSLTLKVKVNVANGEVLYERD 1003  
 QY 902 SAGN-----DOCPSTVTT-PDGCFINTYOPFBDIDVLQVASNE---- 939  
 Db 1004 KAGOLIREVDFTGREICRYDRGRRIATRYPDMHEWY---NEGLVVEQESEWFD 1059  
 QY 940 -----1TQFSNPNVPGALLKAVAEQSLTIVYPSGRKLHENINMKMOSYLTWRG 992  
 Db 1060 EQBSRCLSTPAQSYN-ARQOLIKATNPSPSUVFEDDGCRLCSERINE-QEIVHOM---- 1113  
 QY 993 LENGTHDUTGTIQKI-SDTH-----GRVQTKDOSSTKTTNDDLARHQSQVTDAT 1045  
 Db 1114 --NEADVTLALTRFGERBLHYARGALGBTLSQ-----VNOHAPIQFS--- 1154  
 QY 1046 GHMLTTTVEFDGLANREICRK---LCDBSGHT----LIOQSWLKTQOCANRIVKNGVTL 1097  
 Db 1155 -----YNAVQGEYIQLRSRAGFVNNSHYTATGLLHQHAGRGTEQF-----LQSIQ 1199  
 QY 1098 QRTEQ-----YSDSRNLNQYKCDGARCPCTDVGHSIVTONTYDVGNTACH 1147  
 Db 1200 AHPOOPPRFCTDVHRSYQD-----RAYHVGIE-DDDRWQT---RYHYNANDOITHQ 1248  
 QY 1148 TTFADEGTDHATKFAINTDPCOLTEVHHTPD--MDNR-----LYKDAGRWN 1198  
 Db 1249 YSPQWGNDE-KPQYDNIN--ITE-HLTTPSSMVSIDAQGMLQFOQOOGAVYRR 1303  
 QY 1199 TDNGHNTENFTYDPLGRQ-----NQGQSV-----YGYDP 1228  
 Db 1304 YTAKG-YQDHYDNGRJAKKIYHTRGRPREWYLMTONQITACFPKGDCWHITDA 1362  
 QY 1229 LNLILVUSQKUTDCELYRRETMLVNEVRNGEMTRLRGETI-----AQQ- 1274  
 Db 1363 FGRLIS-KKTKVDSLADLPFLPQIKRKITWRYLWGGDQOLIEETPYADGTIANRACQV 1421  
 QY 1275 ---RASKVLTGTDQSOQSV-ILTSQDKNLSQDQEASAYG-----KH 1310  
 Db 1422 QWLYQPGELTPARYQOKPLHVYVWTDQGTPRIFSECGQASWGRQMLTWGQMFWRYRD 1481  
 QY 1311 KSTANDAST---LGYNCRERADPVSGVTHLGNSYRSYDPTMFHTPSLSPRG-AGCIN 1365  
 Db 1482 GKAENDPNTVTECPFRFAGQYEDBSGLEY-NRFRYIIRBTGQYLSPLSPG-AGCIN 1536  
 QY 1366 PYSY---CIGDPINRSDPSGHLSQAWTGIGMIGAG---LLITAT 1405  
 Db 1537 PYGYVHC---PTGWDVDP-----FGLAGDCDCELLVLIAT 1566

RESULT 12  
 US-08-551-557B-61  
 ; Sequence 51, Application US/08851567B  
 ; Patent No. 652484  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ensign, Jerald C  
 ; APPLICANT: Bowen, David J  
 ; APPLICANT: Petell, James  
 ; APPLICANT: Ratig, Raymond  
 ; APPLICANT: Schoonover, Sue  
 ; APPLICANT: ffrench-Constant, Richard  
 ; APPLICANT: Rocheleau, Thomas A.



FILE REFERENCE: GTC99-03PA  
 CURRENT APPLICATION NUMBER: US/09/328,352  
 CURRENT FILING DATE: 1999-06-04  
 NUMBER OF SEQ ID NOS: 8252  
 SEQ ID NO 5545  
 LENGTH: 804  
 TYPE: PRT  
 ORGANISM: *Acinetobacter baumannii*  
 US-09-328-322-545

Query Match 2.7%; Score 235.5; DB 2; Length 804;  
 Best Local Similarity 20.8%; Pred. No. 4,4e-09;  
 Matches 147; Conservative 92; Mismatches 253; Indels 215; Gaps 30;

Qy 699 YTIHTQQDEQGRHTSTYSDYLNGRQTPDPKVHLSMSKSYDNWGQIANTWS---- 753  
 Db 21 FTQLTTEYDAGATRSNTOAGRVTQKYKT------VSFSTCL---GNAFLROWSKEVSC 71

Qy 754 --YGVSEKITVDPITLTKOLOSNSNNVQTKEVTTYPSQO--PQQTDEBAGHQ 808  
 Db 72 LSNTVTLGDORHVGIEV-----DVNGKEITOPFAGYGRGIPQVKANGA--- 116

Qy 809 SCHILTRDQWDVRKETDAIGQCTIYQYQNYNRVIOITLPDGTIVNRYKAPFSTDUTTD 868  
 Db 117 -TERINIVDDFGNITQHTDADGIVSRKQIDDAAGRIL---YIDTFIVGLANISTFTYDGLTVS 171

Qy 869 IRVNG-----IS-----LGQPTFD-----GOFIH--- 921  
 Db 172 RVVTFGGQLSLRIBKYNGDGLLSSBEDKLINSNKTSINSNKYDAEFLNLFKSNPGPSAITSGT 231

Qy 891 -----ODGGRWATTYSAGNDQCPSTVITD-----GOFIH--- 920  
 Db 232 TSSYDVFDRPTIVNDGTSVTCYQSCSGKTAGVQVTDSPGTTESNLAAQDFSDAILT 291

Qy 922 ------YQOPELDDAVIAQO--VASNEITOQFSY-----NPVYGALKA 956  
 Db 292 LVARKGITGSVVFQTTERENALIJKPKVAVGSSSTQSYTNSNTLAEBKNSISGQTK 351

Qy 957 VARGQSLSPIVYPSGRLKMENTINDMKMKNVWTLRLGLENGY--DLTGFRIOKUSR---- 1009  
 Db 352 YDFTGRITSITHPDPSSVETIKYFOLKDITASRTRWREVERTYSISLAGRKTTTANISEA 411

Qy 1010 --DTHGRV-----TQIKOSIKTTIANDDLARHGSQVTDL-----A 1044  
 Db 412 FDLDTYGRVISHOTKINANDNTNSVWVYNGDNL---QWTSIQVPGNGSVNLNSNQNA 466

Qy 1045 TGHMLTTVEFDGLNREIGRKLCOSSGHLDQOSWLMKIQJANRIVKLNQMR--TE 1101  
 Db 467 LGVFTSIPIVQIQSINTNARQQLTVQAMTDLMSYTMDGLANNISAS--LEKCVLAN 524

Qy 1102 QYSYDSRNLNQYKCDGABCPTDKYGHISVTVNFTYDGNITIACHTFADGSTE-DHATE 1160  
 Db 525 DGYGDKLNRVNUKL-----SDKCGS--VYNTATDYG----TQIMSTVELDQARY 567

Qy 1161 KFANPTDCQLTETVHTRDMPNIRK-----YDKAGRGINITDNHENTPFT 1210  
 Db 568 QSYNNND--DIRKV-----NITSKSSTVAPAIYTYVNVNTSLASVSGSTVSPY 616

Qy 1211 DTWGRLOQGQGSVYGYDPLARLVSQKTDLCEBLYYRETMALNEVR 1257  
 Db 617 DAMGNVNTDVGVRVLTIVDYSRANKGNTY--LYNADGLRVRAVD 660

RESULT 14

US-09-252-991A-31032

; Sequence 31032, Application US/09252991A

; Patient No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenstein et al.

; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 10719\_136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; SEQ ID NO 31032

; LENGTH: 1128

; TYPE: PRT

; ORGANISM: *Pseudomonas aeruginosa*

US-09-252-991A-31032

Query Match 2.5%; Score 223; DB 2; Length 1128;  
 Best Local Similarity 19.6%; Pred. No. 7, 5e-08;  
 Matches 258; Conservative 151; Mismatches 467; Indels 442; Gaps 68;

Qy 225 NFSIGNBENPLTWFSRGYMP---IGKNGLIGQWNTSMWAPGSLKETVNNVNINNOHH--FP 288  
 Db 63 DFALAAPPTPTFARGYLSSNPRIGRLG-RGWWLPGESHLSEDACVLUVDAGRRIGFP 121

Qy 299 QSANLPVLPVYUMLQVPGAGQPAIQAEYSYSHNTVUGGSNGIWNWKLNDYGLMTYN 348  
 Db 122 ALA-----PGA-----ORTS-----GSEBLW----- 137

Qy 349 YGSTBSRSPYKDXEKGHDQVTRIBTYNNHILLTSBECKQONGYQTTETAYA1IGHNFDQ 408  
 Db 138 -----LRGGSSCGRQAQWR-----GMAAVPQE-----LQTBESSVLVLSGHSY--- 177

Qy 409 PSOPOLPLPKTETWRSANSYRSBIT-----ETTPDESBNPLTKVKK-----TQ 455  
 Db 178 -LHFQ-----RCPGIWIWLQASFRAGRTRIEFWRKSGRGLTGYDGRDSAGRSYALWQ 227

Qy 456 KIISPSTHMEVYPPAGEVDNCPEPYCFTRFKITOPVDSBKO--DPEK----FIQ 508  
 Db 228 QACBPS-----ECDD-----GRLP--GVLASHDGPYPDYDQSPGDLWV 269

Qy 509 YRSLIGSQSHVTLKISERHYSATOLL-----NSTLFOINTD-KSBLGILK 554  
 Db 270 YQPSGSD-----LIAVRDQGVVRFAWREHMLVAHGPGELEVREWDVHAPHGKV 325

Qy 555 QTCTKGFKGKTYSVWHPYTYKQDQHQSHISITHDINFTHRSQVSRSYRGRLESITD 614  
 Db 326 QLEAG-----GLTRTFYLER--DATEVDSLGLRVERYEPAGEGG-QRRWMLVRAIGS 375

Qy 615 TKDPTSYKDGKLRJLRTLNSGTPVANTLTYDYEINLQDNRPIPFTTDVQNL 674  
 Db 376 R----SERDYDLFGRL-----VAKRDPLGET 398

Qy 675 RNFEDGAGHRHSQCOLKSDGDKGKFYIHTQYDBQGRHHTSTYSDYLNGRQTPDPK-- 732  
 Db 399 RARRBDGCRMLB--ESPGKARY-----RKRVDETG 429

Qy 733 -VHLSMSKSYDNQGQIANTHWSYGVSKRITVDPITLTKOLOSNSNNVQ-TGKETY 789  
 Db 430 LVELE-----DAMQRWIFERDR-----GNATVVRGPAGSTRAY 465

Qy 790 TSQSQPIOT-LFDEAGHFLQSOCHLTTRGMDR--VTKRTDAIGQCTIYQYONTNRVQI 845  
 Db 466 EDRPLPDRPTIVDPRGER-----RUEWNRFLGLLAUTDCSGQWVRYDYNENGRLVAS 519

Qy 846 TUDGTIVKAPFSTUTLTDIVYQGISLGQTFQSLRSITQSDG-----G 894  
 Db 520 SPPLGQWRRYDPLG--OLIGBLABGSALSYE-YDAGRGRTRADEGHATLFSWGHG 576

Qy 895 RWAYTTSAG-----NDQCPSTVITPDGFQHYQQPE-LDDAVTQVASNEITOQFSY 946  
 Db 577 DLLARVSDAGGERBLSYLHDEAGRIVLAVTNENQVQAQFRYDILDRUYETGFDGRORVRY 636

Qy 947 NPVTGALIKAVAGQSLPTVYPSGRLKMENTINDMKM-----SYWLTG----- 992

Db 637 NADELARREDADGRETTAYVDRDGLRSLRVPATERPALVERYWRWLAQDGLASAGGA 696

Qy 993 LENGYT-DLTGTFQKUSRDTH--GRVTCIKDS---SIFTINYDILNARHGSQVTDL-- 1043



QY 1098 QRTEQSYDSRNLNQYK---CDGABCPTDKYGHSTVTONFT----- 1136  
 Db 1621 ---NRSTEVKONALADDIAPIAIPNPIENIULDINGQNTAQAEANTQIEVKNAVG 1669  
 QY 1137 ---.YDIYGNITA-CHTTFAUTERHATKEPAN-----PDPCLTEVHTRPOMPD 1183  
 Db 1670 BIVESGYVDAAGNVSGLIQVYLGEB--LTFFVVDRAGNRSTEVKONALADDIAPNPIE 1727  
 QY 1184 NIRLUKYDKAGRVINIDNHGNTENFT---YPTLGRLOQGSGVYGPPLNRLVSQKDT 1239  
 Db 1728 NILL-----DANGONFTAOREANTOIEVKNAAGSFIG-----SGSTD 1765  
 QY 1240 L-DCELYKRETMVLNEVRNGEMIRL---RTGETIAQORASKULLTGTDQSOSVILTS 1295  
 Db 1766 MGIVNSGYFVQYL---HGELTFUVVDRAGNRSTEVKONALADDIAPIALIFEN 1820  
 QY 1295 KONLSQEA-YSAVCKHKSTANDASILGNGERADPVSG-----VTHLGN 1338  
 Db 1821 GONFTAQABANKSVKVERVKAJGEVVGSGY-VDSAENVSGYLNOVYLKGEBLTFFVVIDQAGN 1879  
 QY 1339 GMSYDPTLMRHPPDSLSPFGAGGI-----NPVSYCIGDP---INRSDPSHL---- 1384  
 Db 1880 -.RSIEVKOTAF-LDNTAPENATNLVFSEDGSYLSGMKAEPNATIOFDQYGBLNNWN 1935  
 QY 1385 -.SWQAWTOIGMG--IAGLILITIATGGMAMAAAGGIAAIASTSTALAFGALSVTSD 1439  
 Db 1936 NVNWGDTRNIYLNNSNYMIGEVFKV---VVWDQAGNUSEGV----- 1973  
 QY 1440 TSVGALEBDASPKASTIGGWVSMGMAAGLSEAIIKGCTKLATHGAFABGENALIKS 1499  
 Db 1974 --VKAPLDDIAPIAASDLVFNEDGSSLSLGVABP-----NTPIQIFDQNQQ-MNT 2020  
 QY 1500 TSSESRKIMGVTRSLDRELVRNFEQVQIKDHSGYTDIFMGKGEQALIVHGDGFLYT 1559  
 Db 2021 WSOSVNADGCTFTIFGTTNLHGSEFTVI-----VKOLAGNVEASV----- 2062  
 QY 1560 BGKHKHKGKPYTRHTPEQQLYDYLKONIVDLOCGD--KPVHLLSCYGKSSG 1609  
 Db 2063 -----KAPLDDIAPNPIKNTVFDANGOSFTAQAEANSQIEIFDSFGSQIG 2107

Search completed: January 30, 2006, 09:50:44  
 Job time : 46.512 sec

THIS PAGE LEFT BLANK

GenCore version 5.1.6  
 Copyright (c) 1993 - 2006 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 30, 2006, 09:43:19 ; Search time 104.198 seconds  
 Sequence: 1 VVYIKPLKURRITMSDNHFP.....PRKILGSRTEKTVPKTFRP 1673  
 Million cell updates/sec: 6708.668

Title: US-09-889-874A-23

Perfect score: 8879

Sequence: 1 VVYIKPLKURRITMSDNHFP.....PRKILGSRTEKTVPKTFRP 1673

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqB, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_Main:\*

1: /cgn2\_6/ptodata1/pupaa/IS07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata1/pupaa/IS08\_PUBCOMB.pep:\*

3: /cgn2\_6/ptodata1/pupaa/IS09\_PUBCOMB.pep:\*

4: /cgn2\_6/ptodata1/pupaa/IS10A\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata1/pupaa/IS10B\_PUBCOMB.pep:\*

/cgn2\_6/ptodata1/pupaa/IS11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	932.5	10.5	982	4	US-10-365-742-106
2	517.5	5.8	1317	4	US-10-365-742-106
3	420.5	4.7	2364	4	US-10-156-761-7834
4	395	4.4	820	4	US-10-156-761-7990
5	392	4.4	1250	4	US-10-156-761-7572
6	391.5	4.4	2386	4	US-10-156-761-7751
7	386	4.3	2334	4	US-10-282-122A-46565
8	361	4.1	1110	4	US-10-418-861B-55
9	353.5	4.0	893	4	US-10-232-122A-76214
10	350	3.9	1515	4	US-10-282-122A-47600
11	349	3.9	1397	4	US-10-282-122A-43059
12	340.5	3.8	1411	4	US-10-232-122A-43060
13	339	3.8	1426	3	US-10-912-020-340
14	339	3.8	1426	4	US-10-232-122A-42617
15	339	3.8	1426	5	US-10-771-241-340
16	331.5	3.7	1377	3	US-10-815-241-10384
17	331.5	3.7	1377	4	US-10-287-274-467
18	331.5	3.7	1377	4	US-10-282-122A-42731
19	328	3.7	1572	4	US-10-282-122A-69115
20	322	3.6	1383	4	US-10-282-122A-68342
21	317.5	3.6	1565	4	US-10-282-122A-45558
22	314	3.5	1395	4	US-10-282-122A-50586
23	307.5	3.5	932	4	US-10-282-122A-44617
24	306.5	3.5	1530	4	US-10-282-122A-68175
25	288.5	3.4	1531	4	US-10-282-122A-49308
26	297	3.3	1512	4	US-10-282-122A-78386
27	291.5	3.3	1253	4	US-10-282-122A-68914

## ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	932.5	10.5	982	4	US-10-365-742-106
2	517.5	5.8	1317	4	US-10-365-742-106
3	420.5	4.7	2364	4	US-10-156-761-7834
4	395	4.4	820	4	US-10-156-761-7990
5	392	4.4	1250	4	US-10-156-761-7572
6	391.5	4.4	2386	4	US-10-156-761-7751
7	386	4.3	2334	4	US-10-282-122A-46565
8	361	4.1	1110	4	US-10-418-861B-55
9	353.5	4.0	893	4	US-10-232-122A-76214
10	350	3.9	1515	4	US-10-282-122A-47600
11	349	3.9	1397	4	US-10-282-122A-43059
12	340.5	3.8	1411	4	US-10-232-122A-43060
13	339	3.8	1426	3	US-10-912-020-340
14	339	3.8	1426	4	US-10-232-122A-42617
15	339	3.8	1426	5	US-10-771-241-340
16	331.5	3.7	1377	3	US-10-815-241-10384
17	331.5	3.7	1377	4	US-10-287-274-467
18	331.5	3.7	1377	4	US-10-282-122A-42731
19	328	3.7	1572	4	US-10-282-122A-69115
20	322	3.6	1383	4	US-10-282-122A-68342
21	317.5	3.6	1565	4	US-10-282-122A-45558
22	314	3.5	1395	4	US-10-282-122A-50586
23	307.5	3.5	932	4	US-10-282-122A-44617
24	306.5	3.5	1530	4	US-10-282-122A-68175
25	288.5	3.4	1531	4	US-10-282-122A-49308
26	297	3.3	1512	4	US-10-282-122A-78386
27	291.5	3.3	1253	4	US-10-282-122A-68914

Result No.	Score	Query Match	Length	DB ID	Description
1	932.5	10.5	982	4	US-10-365-742-106
2	517.5	5.8	1317	4	US-10-365-742-106
3	420.5	4.7	2364	4	US-10-156-761-7834
4	395	4.4	820	4	US-10-156-761-7990
5	392	4.4	1250	4	US-10-156-761-7572
6	391.5	4.4	2386	4	US-10-156-761-7751
7	386	4.3	2334	4	US-10-282-122A-46565
8	361	4.1	1110	4	US-10-418-861B-55
9	353.5	4.0	893	4	US-10-232-122A-76214
10	350	3.9	1515	4	US-10-282-122A-47600
11	349	3.9	1397	4	US-10-282-122A-43059
12	340.5	3.8	1411	4	US-10-232-122A-43060
13	339	3.8	1426	3	US-10-912-020-340
14	339	3.8	1426	4	US-10-232-122A-42617
15	339	3.8	1426	5	US-10-771-241-340
16	331.5	3.7	1377	3	US-10-815-241-10384
17	331.5	3.7	1377	4	US-10-287-274-467
18	331.5	3.7	1377	4	US-10-282-122A-42731
19	328	3.7	1572	4	US-10-282-122A-69115
20	322	3.6	1383	4	US-10-282-122A-68342
21	317.5	3.6	1565	4	US-10-282-122A-45558
22	314	3.5	1395	4	US-10-282-122A-50586
23	307.5	3.5	932	4	US-10-282-122A-44617
24	306.5	3.5	1530	4	US-10-282-122A-68175
25	288.5	3.4	1531	4	US-10-282-122A-49308
26	297	3.3	1512	4	US-10-282-122A-78386
27	291.5	3.3	1253	4	US-10-282-122A-68914



Qy 933 LQVASNETTQQFSYNPVGALLKAVAEQGSITPTTYYSSGRKLMENINDMKGMSYLWTLRG 992  
 Db 826 GRTPARK-----YDAVGNKISSSTDEGKTTSTGYDVLARLTKVSYPPDQKSY----- 873  
 Qy 993 LENGYTDLTGTIKISRDRHGRVTOQKOSKIKTLYNVDLNRRH-----GSQV--TDLA 1044  
 Db 874 -----NYDAVGNRUTMKGSHGTAYKDNLRLISVANPDGOKSYTNK 918  
 Qy 1045 TSGML-----TTVDPGCLRIGKLCDSGHTLD--IQSMLKTQLANRIVKL 1093  
 Db 919 VSNRVRKMTVPGKTSYSYDVAUNLIG--VSDGHTTSYSWAKNGNLKTMNTPNGV-- 973  
 Qy 1094 NGVLQRTEQYSYNSRNRLAQYKCGAECPTDKYGHISIVTQNFTDIENITACHTFADG 1153  
 Db 974 -----KTB-YSTYQKANRL-----VELINKNTIQVUSYKYLDAAGNRUKDEQLAEG 1020  
 Qy 1154 TED-----HATFKPANPFDPCOLTEVHHTHPMPDN--IRLKDGKGRVINT- 1199  
 Db 1021 VEGSDSELEKESOLITTYG--DKLYRJTKV----DYPNSKTVSYKDSMRNISMT 1072  
 Qy 1200 -DNHGNTENPTDYLGLQNGQSYVGDPLRNLVSQKTDCELY-YRETMVMEVR 1256  
 Db 1073 NVUGIGSTISYKYDAADGULLOSGNTSYSDKGHLKGRVNSTOFMSYSYDB--ANRLK 1129  
 Qy 1257 N-----GEMIRLRT--GETIAQRASKVLTGTSQSQSVLT 1293  
 Db 1130 NWBPFVSNWTPKSSYNFEDGDNRKTITNGE---NAQSTKVLIDINSALPQVITE 1185  
 Qy 1294 SDKONL-----SOEA-----YSAVGHKSTANDASIL----- 1320  
 Db 1186 SDIKKNTCTYGTDLISMNSTENAEVYHYDGGSVRSLSKGKIKOTYLAFAQVQK 1245  
 Qy 1321 -----GYNERRADPVSGVTHLGNGYRSVPTMLRFTPLS--PPGAGGINPSY 1369  
 Db 1246 EIGTFVNDNFRFTOBMDDTGTLYLRARY--YDPSVGRPITKOVIGRVRTOSINRYV 1303  
 Qy 1370 CLGDPINNSDPSGH 1383  
 Db 1304 TTNNPVNUVLDTGY 1317

RESULT 3

US-10-156-761-7834  
 Sequence 7834, Application US/10156761  
 Publication No. US20030119018A1  
 GENERAL INFORMATION:  
 / APPLICANT: OMURA, SATOSHI  
 / APPLICANT: IKEDA, HARUO  
 / APPLICANT: ISHIKAWA, JUN  
 / APPLICANT: HONIKAWA, HIROSHI  
 / APPLICANT: SHIBA, TADAYOSHI  
 / APPLICANT: SAKAKI, YOSHIOKI  
 / APPLICANT: HATTORI, MASAHIRO  
 / TUTOR OF INVENTION: NOVEL POLYNUCLEOTIDES  
 FILE REFERENCE: 249-262  
 CURRENT APPLICATION NUMBER: US/10/156-761  
 CURRENT FILING DATE: 2002-05-29  
 PRIOR APPLICATION NUMBER: JP 2001-204089  
 PRIOR FILING DATE: 2001-05-30  
 PRIOR APPLICATION NUMBER: JP 2001-272697  
 PRIOR FILING DATE: 2001-08-02  
 NUMBER OF SEQ ID NOS: 15109  
 SEQ ID NO 7834  
 LENGTH: 2364  
 TYPE: PR  
 ORGANISM: Streptomyces avermitilis  
 US-10-156-761-7834

Query Match 4.7%; Score 420.5; DB 4; Length 2364;  
 Best Local Similarity 21.6%; Pred. No. 2.4e-19; Indels 423; Gaps 73;  
 Matches 320; Conservative 190; Mismatches 550;

Qy 134 RIHKSGDIEVLTGFNNNAFLDKPKKULNPAGHAIYIDWNFBATQPRLNRYDLDQHD 193

Db 702 RKTGDPGDIVLVINERFTDAD-----NKVGSAA--ELNF----PRIKEHGD---- 744  
 Qy 194 IPLANLBQGLIKILITLPFGQKEGYRTELFLNRLQNSIHNSLGNBNPLTWSGFTPI 253  
 Db 745 -----GATTIVSKOFAN-----ACDIDHPRQS-----NTQDCWQ-KTRPE 781  
 Qy 254 GKGNGILGOMI-----TSMTPGGKLTETVNS-NNNOGHHRPOSANLPUVLP-- 297  
 Db 782 GEDSKTGWFKKFPLVTQVQDPTVTTNDGAPMTISTYEDGAGWHR--TNDPLKDE 838  
 Qy 298 -----YVILMKQV-PGAGOPALOEYSTISHVUGGNGIMWKDNLQYGLMEY 347  
 Db 839 DESWTDWGRQYEVQVTTGAGOK-----TKKSWLYRGSGRTSKAD----- 881  
 Qy 348 NYGSTESTRSYKDKEGHDOQIVRERTYNN-----HILTBESCKOQNGYIQTETAYA 399  
 Db 882 -ASRATKTVDDDG-----NYTDSWSLSSRLSILSRLDQG-TSHERTHJK 927  
 Qy 883 TGEH-----FDPSOPSOFOPLPKTKEPTWSSADNSYRSETETTDES-----GNPLTRVI 449  
 Db 928 YWDHNTAQYDGLPDAFREKEETTNTKVSSGNGTVEETDTEGASTTGFGLPMR-- 984  
 Qy 450 KDKTKQK1SPSHMEYRPAGEVDN-CPPRGFTTRVKKIQTPTUSEFKDFPDKPTEQ 508  
 Db 915 TDDWGSVS-----DNRC-----TTYGRAYNT----- 1008  
 Qy 509 YRSLSIGSOSHVTULKIBERHYS-----ATOLLNSTLFQYN----- 543  
 Db 1019 -NYDSTGAQRWTLQDQWKRHYSGCSSLADSKGQYASLYNATSUDVANKPUDGNPES 1067  
 Qy 544 -----TDKSGRBLRUKQTE-----CTKGENGKTYSVVHKFTYTKODDTIQQ 584  
 Db 1068 RTTKSGSYRSTWGSYDAGRVMWSEPSOSHNRILTYSKSPANTWPLNGVIVTTPDPDGALP 1127  
 Qy 515 SHSITTHONFTIHSOVSRSYTCRLESTPDKTIVTONSYDKLGRLL----- 631  
 Db 1128 AHFALTSTPAWT-----SRFWGKPTSIODANGNVTKMSDAGRLVWWRPTETGSSPS 1180  
 Qy 632 -----TFTLNGTPVANTLYDYEANLQDDNRPPPV 664  
 Db 1181 MKESYTIPTSTNSAGVPAVDGYPHVATHVLSHATVDGGRAR-ETOTENGN 1239  
 Qy 655 TTVDUNGQNLNRFEDGAGRHVSOCCKLPSGDGDKPYTINTQDDEGR-----HH 713  
 Db 1240 GVDATAGNEVPN-----RQSVTRYDAGN--VIGTSAFRNQGTTGGGGSSAKVED 1290  
 Qy 714 TSTVSDVJN--GROQDIPDKHLSMSYSDWQGQIANTHWSYGVSEKITVPTILATK 771  
 Db 1291 LPSYTDVLWDGRATTS---RLQVNGASQDAGRVDPTYDGFSTVWKVDAADTYTD 1346  
 Qy 772 QLOSNNSNNVQNTGKEVTT-----TYPSPQOPIQITLDFEAGHLOSCHTLTRDGMWDRVRKETD 826  
 Db 1347 WQGQSKVYKVEHTCSATVYTAYTVAQBLIKIT-DERFG-DSYTYWAGORTATD 1403  
 Qy 827 AIGCQTIVQYDNTNVRQDITLPG-TVNRKTAFFSTDLTIDIRVNGISLGQOTFDGILS 885  
 Db 1404 AGVSST--BYDKRQISKITSNSNGKTTDIDALGRKTA--VRSGETELAATWDGILN 1458  
 Qy 816 RLTSQDPSGGRWAVTYSQGNDPCTVTPDOSFIHQYOP-----ELDDAVLQASNEI 940  
 Db 1459 -IT---GCK-GQTKTSTSRTDNTNTYTKTBF-DYGRPLKTTVTIPDTVKGLAGS-Y 1510  
 Qy 941 TQPSVSYNVTGALLKAVAEQGSITPTTYYSSGRKLMENINDMKGMSYLWTLRGLENLY-- 997  
 Db 1511 TGTGTYD-----AADRTSVAAPAAGLPAETV-----ATYDGGGR 1547  
 Qy 998 -----TDITGTDIYKISRDRHGRVTOIK-----DSIKITLNVDLN--RHGSQV 1040  
 Db 1518 PLRQSLAQMTVIRSTGDAYDULTRDSGVDSLPGIGAAQRTYSTDSNCIRELVA 1607  
 Qy 1041 TDLATGHLTTTVERFDGNNREIGRKCLDSSGHTLDDOCSWLMQQLANRIVKL--NGV 1096

Db 1608 T-----TTT-----LNKVNSERQKDT-----YTYVLAGKTELEREQASQ 1642  
 QY :  
 Db 1097 LQRTEQYSVDSRNL-NQIK-----CGAECPTDKGHSTVQNFTYDIGNITACHT 1148  
 QY :  
 Db 1643 TADSOCFLYDDQARLTNAVTTHRGICADKTKTASDFKGKTAQOTAYTVDRLGNTSIN 1702  
 QY :  
 Db 1149 TFDAGT-----EDHATKEFANPPTDPCOLTEW-HHTHPDMPNIRLKIDKAGRVI 1196  
 QY :  
 Db 1703 TDSAGTATLHDLYPGYDGTWTWITANANQPHGVRKIKHKGSPITKSQDLYFDDGMK 1762  
 QY :  
 QY 1197 NITDNHGNENTFTDGLQO---NGCOSV---YGDPP---LNRLVSKT---DTL 1240  
 Db 1763 QRVE-PGTTDWTPOGQLEAVKITKSSGELTRYVADAGNVLVRTPOETASIDM 1821  
 QY :  
 Db 1241 DCELYRETMVLNVNE-RNGEMMIRLRTGTTIACQRASKVLTUGTDSPSVILTSRDKN- 1298  
 QY :  
 Db 1822 ELRITNGTWTWTRVYASGTTAVAMRTTGTATN--GKTVLWMDTQASTOLAVDASTG 1879  
 QY :  
 Db 1299 -LSQBAYSAVGKHK-S-TANDASILQYNGERADPVSGVTHLGNGYRSYOPTMLRHTPSL 1356  
 QY :  
 Db 1880 ASTRRYTPFGDENSGLPTGTDNGFLKGTEPDSTGSLL-GARAYDPNLGRPLSPDPL 1937  
 QY :  
 Db 1357 -SPFGAGGINPYSYCLGDPTRSDPSGHTSWQANTGIGNGIA 1398  
 Db 1938 ATPYAPQNLYNAYSVATNNFISYSDPSG-LCRRDICGDGYPVAG 1979  
 QY :  
 Db US-10-156-761-7990  
 ; Sequence 7990, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHITAKI  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 1509  
 ; SEQ ID NO: 7990  
 ; LENGTH: 820  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces avermitilis  
 ; US-10-156-761-7990

RESULT 4  
 US-10-156-761-7990  
 ; Sequence 7990, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHITAKI  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 1509  
 ; SEQ ID NO: 7990  
 ; LENGTH: 820  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces avermitilis  
 ; US-10-156-761-7990

Query Match 4.4%; Score 395; DB 4; Length 820;  
 Best Local Similarity 23.8%; Pred. No. 2.7e-18; US.10-156-761-7990  
 Matches 242; Conservative 375; Indels 280; Gaps 50;  
 QY :  
 Db 501 DDEPEKFIQYSLIGSQSHTLKIBERRHYSITQLNSTLQQYNTDKSELGRLLQTEC-T 559  
 QY :  
 Db 26 DTEGKTFVFRYDAWGRVKIT-----TAEGRTVFTYD-DANRVTMSLRGTGENS 74  
 QY :  
 Db 560 KGEGNKTYSVVHKFTYKQDDTLQSHSITHDNFTIIRHSQVRSRYTGRIFSDTDTKDIV 619  
 QY :  
 Db 75 DGHGTFTW----TAYTSLSVTAGITTRDPTP-HATKXQHSDGQVSDVDAKGK 127  
 QY :  
 Db 620 TOMSYKLGRILRT-LNSCTPYAATLTDYEL-NNLQDNDRPPFVITTDVNGNQLRN 676  
 QY :  
 Db 128 RSTKDANHSIDTSTDAMGSGTTPGNTDVGCFNTRNLN----TTQPTGKTN 178  
 QY :  
 Db 677 EFDG-AGRHVSQCLKSDGQKFYHTQYDQGRHISTYSDLVINGRQQTUPDKVHL 735  
 QY :  
 Db 179 HMQTAGADVPKDSTNPDGERTDFT----YDAVN-----TM 211

RESULT 5  
 US-10-156-761-7572  
 ; Sequence 7572, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHITAKI  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697

QY :  
 Db 736 SWSKSYDWMQQTANTHNSYGVSEKITVDPITLA--TKQLOSNNSNNVOTGKVTYTPS 792  
 QY :  
 Db 212 SWAQTGREGGNVSYT-----NPASPTCGGGFQRCRTEKTOIAKCVTT---- 256  
 QY :  
 Db 793 QDPIORTIFDRAIGHLQSCH-----TLTRGWDWDRYRKETTAIGCITYQDYNM-RV 842  
 QY :  
 Db 257 -----TFRYDSAGNLDTWTTPAPLAKTTTYDAGRTRKVTDARGVTWTTYDNRDRTRI 311  
 QY :  
 Db 843 IQITLP-----DPIIVNRKYAPSTTDITDIRVNGISLGQQTPFDGLRLT-QSQ 891  
 QY :  
 Db 312 VDTINNKARVEYWTGDDENSLTOR--TSTGTRKD-----PDLORETINTLQ 356  
 QY :  
 Db 892 DGSR-WMAYTYSAGNDQCPSTVTPDGOPIHQYQPBDDAVLQVASNBITOOPSIVNT 950  
 QY :  
 Db 357 DGSQTLIAY-----TPEGTIVYQ---DPAG-----TVDYTMWEE- 388  
 QY :  
 Db 931 GALIKAVAEQSLSTPIYYPSGRK-MENINDKKMOSYLTWLGLENCGYTDLGTONIS 1008  
 QY :  
 Db 389 -----NKLAELKDOPAGRVTYKVNNDVRTT-TVPG-----GTVQKVD 426  
 QY :  
 Db 1009 RDTGRVYQVIKOSKITT-----LNDDLNRHIGSV--TDLATGHLMTTVEGTL 1058  
 QY :  
 Db 427 PDNSRSPRKT-SPKGPVDPDLYASVGGTSGTTEGKRSRSTSIDNVNTVQG-LKTYTNGA 484  
 QY :  
 Db 1059 NRBIKRKCDSSGHTLDIQQSMKWTQOQIARVNLQVNGULQRTQYSIDSRNRLQVYKCDG 1118  
 QY :  
 Db 485 GRFSYAR-EKKGTLN--SSMUCYDAGNL-----TSQGTAJGCPEATTYTN 531  
 QY :  
 Db 532 AQLTAKNGS--TINNSYDKLQNETAGAST-PEGTRTAKW----TDHSQITSL-- 578  
 QY :  
 Db 1179 PDMPDNTILKYDKAGRVINUNITDHQGNTENFTDGL-RLQOGSVYGDPLNRSQK 1236  
 QY :  
 Db 579 -----TWGK--TYYGTYK--TRLGDTYFHNGPIGLG----- 614  
 QY :  
 Db 1237 TDYLDCELYRETMVLNVENRNGEMIRLRTGTTIACQRASKVLTUGTDSPSVILTSRDK 1296  
 QY :  
 Db 615 TSTAGYDGFGNR-----EPGFTINSWTTGKAYCYLTDALGSVIALTDE 658  
 QY :  
 Db 1297 --QNLQSQBAYSAYVGKHKSTANDASILGNN-GERADPVSGVTHLGNGYRSYDPLMRHT 1352  
 QY :  
 Db 659 TGAKNTVAYSPRGVTRASTSEENQPRTFRAGGYQD-VGTYLHYGARRY--YDNIIGRNS 715  
 QY :  
 Db 1353 PDSLSUPERFAGGINPYSYCLGDPTRSDSGLHSLSQWQAVTGGMGTAGLILITATGGMATA 1412  
 QY :  
 Db 716 PD---PSGEBB-NPYLVAEGLGDPNRIIDNGLFSPSDALDTGSDIFGVGCLAGVTAAB 771  
 QY :  
 Db 1413 AGGIAAAIASTSTALAFGALSVISDITSIISGALEDASPKAISLGMVSMGMGAAG 1469  
 QY :  
 Db 772 TGSIAAYAA-----VCGVAGVGSVAVGVGLAVVGSCALG-GAAG 810

PRIOR FILING DATE: 2001-08-02  
 NUMBER OF SEQ ID NOS: 15109  
 SEQ ID NO 7572  
 LENGTH: 1250  
 TYPE: PRT  
 / ORGANISM: *Streptomyces avermitilis*  
 US-10-156-761-7752

Query Match 4.4%; Score 392; DB 4; Length 1250;  
 Best Local Similarity 21.2%; Pred. No. 8.7e-18; Mismatches 452; Indels 368; Gaps 66;

Matches 271; Conservative 187; Mismatches 452; Indels 368; Gaps 66;

QY 254 GKGTLGQWITSMTAPGGIKETVYNSNNNQGHFPQ-----SANLPVLPYVTLMKQPG 307  
 59 GMEGVAEQGLRK-TARNAVTEDAADASRORGAKICKEPKDVTARMULPQI---DVS 113

QY 308 AGGPAIQARYSYTSHNYVYGGGSNGIWNNKLDNLV-----GLMTENYGSTES 354  
 114 PGOLPLVVKQQFESSYRLGEWGPFTWSSTLDQLEVDTAGVVLVGLVEDGLVLAIPH--PA 170

QY 355 RRYKDEKGHDQIVIERTVNHYHILTSCKQONQYIQTETAYAIGHNFDQSOPSQL 414  
 171 PGVPTPLSPHGRPNLSLDRTPGGTYL----TDQKSHIR----- 203

QY 415 PKTKTETMRSADNSYRSEITTEPDESGNPLTKVIKKCTQKLISPSTHWEYPPAGEVD 474  
 204 -----HFVDRSLAVLLEQLDRNGWIT-----PEY-----DAD 233

QY 475 NCP---PEPYGPTFRVK---KIIQTPYPSFKDPEK-PIQTRYSLIGSQSHVTLKIEE 526  
 234 GAPIYLTSGGYYRURISTEAGRVIAHLHLSAVAGGGDODLIRGY-----TDGHL----- 283

QY 527 RHYSATQLNST--LFQMTNDPKSBLGRJLKQTB-----CTKGENGKTV 569  
 284 -----TEVUNSSGRPLQVCD--ELGRITSWTDFNDHSFSYATDDEDRCTH-QSGAAGTL 335

QY 570 VHKETYTKD-----DTIQOSHISITHDNFTIHSQRSVRSVTGRLESDDTQDVIQM 622  
 336 RSTPAYGAIDPPTGCAHTTWTDSYQOTH--YLINR-----RCQVAAETDALGAVY 386

QY 623 SYDKKGRLITRTLNGSTPYANTLYD-BLNQNODDNRPPFTTDVNGNQLRNEFDGA 681  
 387 QDRYNRLLSQT---DPIGHTISPRYDDAGNLVAATP-----DGREAIAETVAL 433

QY 682 GRAYSOCLKSDSAGSKPYTHOYDEOCHRHTSYDSVLTNRQQTDPDKVILMSKSY 741  
 434 GLPVK--LVNPDG----TITROTFBERG-----NLTJSVTBPS-GOTTRGY 472

QY 742 DNWQOIANTHWSIGYSEKTV-----PITLT---ATKOLQSNNSNNVGTKEVTTYPSQ 793  
 473 DEGCRLTSMWDPPLHHTGIVCDRAGLPLWTPDPLGAIVTRYERDAFGRTA--ITPDIGAT 530

QY 794 OPIOTLTFDEAGHHLQ-----SCHTLTRDGDWRVRKEDAIQGCTIVQYDDNMRVIOIT 846  
 531 TRLEWTV---EGHLSRRTAPDGTSBSWTIDGEGENCTSHIDPPVGCVSLFHYTHFLDLPTART 587

Db 775 --GH-ITLHSFDSLGRUTSQS-LGPAGRTOHQY-----TYPD-----GH 814  
 1129 SIVTON-----FTYDYGNTAACHT-----TF--ADGTEDHATPFKANP----- 1165

QY 815 LIGEDBOSKREFDLILAGRIVAHVNMTETVADAVGNGTSASWAGHOCQAIGR 874

Db 1208 FTWDLGR-----LQCGQSTGVIDELARLUSQKDTLDCLYXPERTMNEVR--N 1257  
 935 YTDPLGRTAKLRLABGCTLV-----ERUDFTWDGTCVLEBOTTWSPDPLPHQWTLWD 988

QY 1258 GEMTRLRLRGTEITAQQQRASKVL-----LTGTSQSVLTSKQNLQOBAYSV 1307  
 989 HQSLRPVQTERWAADRAQEDSRFAIVDVLVGESE--LLDRREIVWRTRALW 1045

QY 1308 GHKUSTANDASI--LGYGBRADPVSYGTHLNGYRSIDPTPLMRFHFDPSLSPPGAG-GI 1364  
 1046 GSTWAKOSTAYPLRFQYQYDPETGJHY-NYFRHYDPETARYVTD--PLGLGPAP 1100

QY 1365 NPVSYCLGDPINPSDPSG 1382  
 1101 NPAAY-VNPNPHMWADPLG 1117

RESULT 6  
 US-10-156-761-7751  
 Sequence 7751, Application US/10156761  
 Publication No. US20030119018A1  
 GENERAL INFORMATION:  
 APPLICANT: OMURA, SATOSHI  
 APPLICANT: IKEDA, HARUO  
 APPLICANT: ISHIKAWA, JUN  
 APPLICANT: HORIKAWA, HIROSHI  
 APPLICANT: SHIBA, TADAYOSHI  
 APPLICANT: SAKAKI, YOSHINYUKI  
 APPLICANT: HAYTORI, MASAHIRA  
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 CURRENT APPLICATION NUMBER: US/10/156, 761  
 CURRENT FILING DATE: 2002-05-29  
 PRIOR APPLICATION NUMBER: JP 2001-204099  
 PRIOR FILING DATE: 2001-05-30  
 PRIOR APPLICATION NUMBER: JP 2001-272697  
 PRIOR FILING DATE: 2001-08-02  
 NUMBER OF SEQ ID NOS: 15109  
 SEQ ID NO 7751  
 LENGTH: 2386  
 TYPE: PRT  
 / ORGANISM: *Streptomyces avermitilis*  
 US-10-156-761-7751

Query Match 4.4%; Score 391.5; DB 4; length 2386;  
 Best Local Similarity 20.1%; Pred. No. 2.6e-17; Mismatches 688; Indels 601; Gaps 93;  
 Matches 379; Conservative 214; Mismatches 688; Indels 601; Gaps 93;

QY 907 QCSTSTVTPDQFHYQVOPELDDAVLQVAS--NEITOQPSY-NPVTGALLKAVACGSL 963  
 625 LVABTDF--DNRATLEY---DFAGRLASRNLVALGOMIAFERNELQGIVRKAAGT 677

QY 964 TPIVYPSGLKMEMINDMKKOMSYLWTLRGLENGTDLTIQKTSRDTHGRVQIKUSSI 1023  
 678 TIVADFTQLOQA-----TGPDTAATLITIRDHRGMRSBEAVNGR 716

QY 1024 KTTIANYDDLNTRHI-----GSOVT---DIA-----TGHMLTTVEFDGILNTRIGKLCQD 1068  
 717 ALTYDYEDELGRRTRTTPSSATTIWSYDVAGRRTSMGSGS-IDFVYDEAGRELGRF-- 774

QY 1069 SGHTLUDIQQSLKTOQLANRIVKLNGVILQRTQSYSDSRNRLNQYKCDGAECPTDKVH 1128

QY 258 ILGQWITS-MTAPGGIKETVYNSNNNQGHFPQSANLIPVLPYVTLMKQPG-----AGQ 310  
 798 SKTQWPKFELVTKVBUPTV--TINQDG-----APMTTSTYEGGAGWRFQD 844

QY 311 PAIOAB-----YSYTHNYVYGGGSNGIWNNKLDNLG-----MTBENYGSTESRXKD 359

Db 845 PLVKODEDESWSDMRYQYEQETQVITGADTGKTKKKWLYRQLGDRTSKTDTSATKTVND 904  
 Qy 360 KEHDHQIVRERTYNN-----HLLTSECKQONGYIOTTERAYAVALGN---FUSQ 408  
 Db 905 GDG-----NHYADHAWLAHTLTSLRDDTDI--SHERTYHTWSHNTAQDGL 951  
 Qy 409 PSQFOLPKKTEWRSADSNSYRSITETPDES-----GPNPLTKVIKDKKQKILSPS 461  
 Db 952 PDAHFVRESKSTTNWKISGWRHEWVNTEDSERASTTFGLPR--TDDWQSQNSVS-- 1006  
 Qy 462 THW3YYPAGEVDN-CPPSYGFTRPVKKIIQTPYDSEFKDPEKFQIQRYSLGSQHV 520  
 Db 1007 -----DNRC-----TYGRAINT-----NYDSIGAQRWT 1031  
 Qy 521 TLKIEERHYSATOLNINSTLFOYNTD-KBBLGRLLKQTCCTKGENGKTVSVVHKFTYQ 578  
 Db 1032 VWDODVKHYSVG---CSSIADSNQDGYSITSYLDNATSIDANKPVDGNAMEV--RTYK- 1084  
 Qy 579 DDTILOQSHSITTHONFTIHSQVSRSYRTGRFLSPDPTKOVITQMSYDKUGRLTRTNSG 638  
 Db 1085 -----AGRYSSTWG-----YDKAGRIVS--EDG 1107  
 Qy 639 TPYANTLTID---YELNNLQDDN-----RPPFVI-----TTDVNGNQ 673  
 Db 1108 KHNKRSTITYSPANTIPMNGISTSTSPPDGATARGPLTTEWISRFWAPTSKDANGI 1167  
 Qy 674 LRNEFDGAGRHSVOSCLKDSPDG--KP-YTHTOQ-----VDEQGRHRTSY 718  
 Db 1168 TKVTLDAAGRVEVWKPTEGSSPSMKFSYTIPSTNSACVPDSADGYPHVASHLQGS 1227  
 Qy 719 DYLIN-----GR-----QOTPDKWHMSKS---YDNWSGIA----- 748  
 Db 1228 TYLVSHAYTDQGLGRARETQFPLPSDUDPATQIVPFROVAVTRYDSAGQVTGASAVFRNQ 1287  
 Qy 749 NTHWASYGV-----BKITDPILTAKTQLOQNSNSNWNQCKEVNTY-----TP 791  
 Db 1288 GTAGSGGPSSQQPSDLPSTSDDLVWDAGRTWSSQJQVKCPTPKAGRVDPSYLDYTP 1347  
 Qy 792 SQQPQITULFDEAG-----HLCOSHTRDGW---DRVRKETDAIGOCTTYQDYNMR 841  
 Db 1348 VDSTATDPTIVVYQGVSKVVEHTASSAYTAYGTYAKGELAQITDPRGNTLYTDMAQ 1407  
 Qy 842 VIOITLDPGSPITVNRYC---AFSTUTLDIRVNGISLSQQTFCGSLRSLTOSODCG--R 895  
 Db 1408 RKTDDDPDAGLSSSEYENNGQVSQRTATNDQVTLYG--YDNLSRATSVRGADELA 1464  
 Qy 896 WMY-----TYSAGNDQCPSTVTPDGCPIH-QYQPELDDAVLQVQSNBTQO 943  
 Db 1465 AWVMDGDPAACTGGKSGOITSAVSRDASGNTYTTKTKGP-----DERGRPLNTVLP 1514  
 Qy 944 FSYNPVTGALLKAVA--EGOSSLTPIYP-SGRKMENT---NDMKMWSYLWTRGLENG 996  
 Db 1515 TVNGLAGDVTISVYDAAHHTSYPAANGKLAKEKVTTDYQGPTRLSSLG--- 1570  
 Qy 997 YTDLGTGIOKISRDTRGVRQIKOSISKTM-----YDDNRHGSQ-VTDLAC 1046  
 Db 1571 ---GTAVIDNTYDAYGRLV-EQGAEFGGNGIQMQRQYGDSSN--GTRWLRSIAT- 1622  
 Qy 1047 HMLTTVEFOGLNREIGRKLUQDSSHTDLOQSMWLTQQLANRIVKLNGVLRQFQSTD 1106  
 Db 1623 ---TTTINDLUSEAKDTYLUDNTGKTELBR-----QASGTAQSQCRLYD 1666  
 Qy 1107 SRRNLIN-QYK-----CDGABCPTDKYGHISIVTONFTYDIGNITACTHTFADGT--- 1154  
 Db 1667 DQSRUFLAYHTTAGNCADTTKTTSPKGSPYQRCYTYDRLGNLQSVPTNSAGAATR 1726  
 Qy 1155 -----EDHTAEPKAFTPDCOLTEVHHTPDMPNTIRKDK-----AGRVINIT 1199  
 Db 1727 DYLYPGYDDAGCTWTTIAN-----ADOPHGV-KINKVSAGTTAAGITYYA 1771  
 Qy 1200 D-----NHCNTENFTYDGTGRL-----QNGQGSVYGDPLNRLVSQKTPDCEL 1244  
 Db 1772 DGAMKORVEGSTTDTWSRUGLATVKTITKTSGLSDLTRYTYDASGNLLVVRTP----- 1825

RESULT 7  
 US-10-282-122A-46565  
 Sequence 425565, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Karl  
 APPLICANT: Zvokind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trwick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forayth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELTTRA.03A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09

Qy 1245 YYBETMLV---NEVRNGEMIRLRT----GETIAQRA-----SKULLTGDSQS 1289  
 Db 1826 --QETVASSIGGELRTDGDISATATRYSFGATVAMENTDGINTVNGKITYMGDQAS 1883  
 Db 1290 VILRSDKO--NLSQEAVASYGKHS-TANDASTLGYNGERADPVGTVHNGQYRSYPT 1346  
 Qy 1884 TQIAVDAATGATARRYYPPFGDRRSGSPLPTGTMHGFLKTEDINTGLSL--GARANDPS 1941  
 Qy 1347 LMRFHPTPLS-PFGAGCINPSYCYLGPINNSDPSGLLPSGKLMNSDGTGDSGWKKGPTV 2001  
 Db 1942 LGFLPLSPPLSTYDPOUNLSAVSYGNDPINTSDPSGLKLMNSDGTGDSGWKKGPTV 2001  
 Qy 1398 G-----LL-----LTATGSMIAAG-GJAA 1418  
 Db 2002 GGGDTPPAPTVQDQLVDRPREENGMDAQLAQVMWPHYQPTQGSCYWDAVGDSGRT 2061  
 Qy 1419 AIASTSTML-AFGALSVTSSTDTSIV-----SGAILEDASKASSTIG 1459  
 Db 2062 GMACGPRACSEAFRWKETHDFAKAKRVAATPCVENPKRCGADNGAVDSMKEASEAVP 2121  
 Qy 1460 WVSNGMGAAGLABSAIKGTLAHTGFAEDGENALK--SSESSESTIKWG-----V 1510  
 Db 2122 LIAGEMGKA-----FSKVAKRGCSFKPVTIRVLMKGKTPKPLGKPGVLDVADP 2172  
 Qy 1511 TRSLDREI-----VNBECQVIKHS-----RGYTDNFMGKGEQALV 1548  
 Db 2173 TSGHREREVTAVHLNHDVLDSLTRGDLGRITQIHTTARHRIWDDTAQWNEQASGRU 2232  
 Qy 1549 HGDKGDFLYHTEGKHH-----NGKGPYTHTPBOLVYDJKDNNDVLDLQGDKR-- 1597  
 Db 2233 TGHK-----VNTSGNHQHATITSVLAQRGADMVLTVEGLITY-----YVLAGETVU 2280  
 Qy 1598 VHLSICSYKSS-----GRADKQAKYNTIRPVIAINSKNTISQSLARIKRDFFLKSTHS 1651  
 Db 2281 VHNGSCWASSTNRKTSVKHAFGHWKH-----KSEFPNNLNNAKEYVYAGTDLRS--- 2330  
 Qy 1652 YDPRKILGURTEKTVKPKTPR 1673  
 Db 2331 -DPS--VLTTRANGDVIRFNP 2349

PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 46565  
 LENGTH: 2234  
 TYPE: PRT  
 ORGANISM: *Bacillus anthracis*  
 US-10-282-122A-4565

Query Match 4.3%; Score 386; DB 4; Length 2234;  
 Best Local Similarity 18.7%; Pred. No. 5.7e-17;  
 Matches 339; Conservative 236; Mismatches 633; Indels 608; Gaps 76;  
 Qy 21 PTOMANNFSAVSSGVDPR-----TGLUNIQTTLGHIVG-----NGNL----- 57  
 Db 748 YSKVGHAESEBISHBFVPRFQSQDSFGFMEDYWDWASIPVWNGKVNATNGNPFMSBRKDITLG 807  
 Qy 59 -GPTPLPLSISPLANKTIGPGFQPNFGLSV--YDRKNLISLIS-----TGEN- 102  
 Db 808 RGDPSVVERTYNSQSKKGFLGGCGWSSLERBWADGNLNLJISTOCANTPFTRCGDK 867  
 Qy 103 -----YRKEITDKTVKLQQKKUDNLRPEKOLKENCYRITHSGDIEVL 145  
 Db 869 YQARTGIVLEIKQVGGVBEIKKDQTV-----TFYKSGDAQR 905  
 Qy 146 TGTNNNAFDLKV-----KGLLNZAGHALYIDMNFEATOPRLNRYDLDGHDPL 196  
 Db 906 IETYTKDQKIGNTTYEVDGASRLRSKVKVNGKELVL---QYDGHNKAARVI---GPDKT 959  
 Qy 197 LNLEYQGKJLKTILTUPQQ---KGYRTSLRPLNRQLNISHNFSLGNENPLTWSFGY--- 250  
 Db 960 ITFNYGDPLLVSSSTPBGKVQYGD----NGVLTSIYDPOHTAKPYKSYAYEND 1013  
 Qy 251 -----TPIGKNGIL---GOWITSMTARGGLKETVNTSNNOCHHFFPOSANTPVLVPTL 301  
 Db 1014 LVRKTDPIGKATLALWNGSKVLTINPKGRKWTYND----AGNPV---- 1057  
 Qy 302 MKQVPGACQAPAOEKEYSTISHNYVGG-----GSNGIWNWKNUDLNLGLMTEN- 348  
 Db 1058 -KTVEDVGRNLNTTSYEVANNLNVKTTPKNQETATYDGNEVNTSVTDEMGTPEKBNK 1116  
 Qy 349 -TGSTESRRYKDE-----GHDQIVIERTTNTNHLLTSCQKQNYIQTTETAYAI 400  
 Db 1117 DNGITKATDNEDTRTVAVGANTEVSTDQGAN----TSVTHHDQYGNIETSKELS 1171  
 Qy 401 ISGNPDSQPSQFOLPKTKTETMSADNSYRSETETTDFESNPLTKVIKDQKTKITSP 460  
 Db 1172 AGSNLQIOPSP-FEM---NCTEKWVDTNNSGSI-----KDTAPRGGLGGSSLKTK 1223  
 Qy 461 STH---WEVYPPAGEVDNCPPPEPYGFTREVKKII-----OTPDSERFKDP 503  
 Db 1224 ATTNNDWGYIA1QEVTLRPNNTYTLSGNWKVTDLNGAAPPNVQSNENGAGTDDGMHDTR 1283  
 Qy 504 EKPIQYRSLISOSHVTLKIBERHYHSATOLINSTLQYNTKSELGR-LIQTECIGK 562  
 Db 1284 HKVQGTSDWVRQ---VTKTKEQ---TRKVLYLOVENGKASATSAEWFKRQIQLRGE 1337  
 Qy 563 NGKTYSVVHKFPTYKQ---DDTLOQ-SHSITTH-----DMPTIHSQVRSSRTGRLFS 611  
 Db 1338 VSSSFNPVTLNSPREENWPDGFVQWVSCSOICERNPVSDDFTGHSISIWER-----S 1391  
 Qy 612 DDTDKDVTQMSDVKLSGLLTRTLNSGTPYANVITYDVELNNLQDDNRPPFVTTDVG 671

Qy 672 NOLRNFPIDGAGRHSQCKDSDGDKFTIHQYDQEGRRHTSYSDYLNRQQDIP- 730  
 Db 1429 -----APDKLSDKDVAVLABTYQG-----TVMVYTRPSGMDW 1464  
 Qy 731 -----DKVHLSMSKSYDNMCOQIANTHWSYGVSEKITVDPPTLTATKQLOGSNS 777  
 Db 1465 NRSAAVIAKKPPIQKIEFLFLFRKONKG-----KWFDDIRLEGNALIKE 1511  
 Qy 778 NNQTGKXVTTYPSOPIOITFDBAGHLOQSCHLTDGNGNLSLQGQTFDGLSLRQSQDGRVW 837  
 Db 1512 YD-NDGNVATY-----DEBQK---NTFTYDASGNKSETDEBKNTKLYDN 1555  
 Qy 838 NYRNUVIQITLPDCTIVNRYKAPPSTIDTIDIRUNGISLQGQTFDGLSLRQSQDGRVW 897  
 Db 1556 KDNLLTKYTKLNGKTSVNRYD----- 1576  
 Qy 898 AYTYSAGNDQCPSTVITPDQFQHYQOQPELDAVLQVASNETTQOFSYNPVGALLKAV 957  
 Db 1577 ---HNGETTERSVMPFGKQTQHKEYEDVDNQNTVYDALKNRIEN-TYDENANKIKRM 1631  
 Qy 958 ABCSOSLTIYPSGRKLKENINDMKNSYLWTLRGLNGYDUTGTOKISDTHGRYTO 1017  
 Db 1632 PNQSLIESVYDADRVVGEK-----RNGKDSFT-----FERDONGQVTK 1670  
 Qy 1018 IKD---SSKTTIANYDILRHIGSQVTLATGMLMTTVEFDGLNREIGRKLCDSGHTL 1075  
 Db 1671 VKDLVNGVERTKTYDKADR-----VTS-AT-----DSRGCRD 1702  
 Qy 1076 IQOSWL-----KTOQQLANRIVKLNGVQLRQEYQSYDSRNLNQ--YKCDGABCPTYK 1126  
 Db 1703 ---WATDKANSKTEKLEKEQTWTQGYY--TINKVSD---INTLDQNIRVTDSQ----- 1748  
 Qy 1127 GHSLVTONFTYDIGNTACHTFADGFTEDHATPKF---ANP-----TDPQCOLTEVHT 1177  
 Db 1749 -----TYPFDYDQGNY-----RUYTAGKGGSTFNYDOANKLKDLYVGTNSNLLSERYE 1799  
 Qy 1178 HPDMDPVNLRYKQDAG-----RVINITDNGNTENFTYDTLGR-----LQ 1217  
 Db 1800 YDOSGNRTRKIKHGAGGRKVTTENFVYDEINQLINEVNGNTKSYTYDFGNRTSVKIE 1859  
 Qy 1218 NGOCS---VYGDPLNLVLSQKTDLDCBL----- 1244  
 Db 1800 NGKETKSAATNEGNGQVKGREGESLTWVNGNRTSPGKYKWTWNEDQIVAITKQCBN 1919  
 Qy 1245 -----YRTELMLVNEVNGEMIRLLXLTGETI-----IAQO 1274  
 Db 1920 AFATYKDDBEDNERIERKNGQVTRYFQGDSINPLYFETDGNSVLRQVYSAQARLMK 1979  
 Qy 1275 RASKVLTGTDQSOSVI---LTSKDQNIQSQAQSAYCR-----HKSTANDASLIGYNGERA 1327  
 Db 1980 AOCQTYLTHYHNPGRDVAMTNOQDKEVVATYEDAWGNVLTSDTKGIAD-NPFGYAGMY 2038  
 Qy 1328 DPVSGVTLHNGYRSYDFTLMEFHTPD-----SLSRPGAG-----GINPYSCLGDRINR 1377  
 Db 2039 DKEIGMTL-----LARYNPEHGVFLSPDGPEDDPVTMNGYTADNPVMM 2088  
 Qy 1378 SDPSGHLTSQWQMTGIGMIGAGLILITATGGMIAAAGGIAAAIASTSTTALAFGALSUTS 1437  
 Db 2089 TPDGKWA-----LVPVVIAGAWAARRGAKAI-----RYGAKYGGKAWKS 2131  
 Qy 1438 DITSIVGALEDASPDKASILGWWVSMGMAAGLAESAAGGSTKLATHLGAFAEDGENVLL 1497  
 Db 2132 -----GWDY-----GKVKVAKSGCNKGKSIACKPRIMKVGR---I 2163  
 Qy 1498 KTSSESSRKRKNGVTRSLDRERIVNBERGQVQKDHRSGYTDNPMGKGQIAUWHDGKOFY 1557  
 Db 2164 KJDNDKQKGWVYIYSTK-----KRGKRTVSS-----SFEPHTPHNGHG 2203  
 Qy 1558 HTEGNGKNG-KGPYR 1572  
 Db 2204 HQQNKVSKYQCKWR 2219

RESULT 8  
US-10-418-861B-55  
Sequence 55, Application US/10418861B  
Publication No. US20040010131A1  
; GENERAL INFORMATION:  
; APPLICANT: da Silva, Ana Claudia Rabera  
; APPLICANT: Farah, Shaker Chuck  
; APPLICANT: Quaggio, Ronaldo Bento  
; APPLICANT: Reinach, Fernando de Castro  
; APPLICANT: Ferro, Jesus Aparecido  
; APPLICANT: De Oliveira, Julio Cesar Franco  
; APPLICANT: De Laiia, Marcelo Luiz  
; APPLICANT: Setubal, Joao C.  
; APPLICANT: Furjan, Luis Roberto  
; TITLE OF INVENTION: Isolated Xanthomonas nucleic acid molecules, proteins encoded the  
FILE REFERENCE: FAPESP 205\_1 US  
; CURRENT APPLICATION NUMBER: US/10/418,861B  
; CURRENT FILING DATE: 2003-04-17  
; PRIORITY APPLICATION NUMBER: US 60/374,620  
; PRIORITY FILING DATE: 2002-04-22  
; NUMBER OF SEQ ID NOS: 85  
; SEQ ID NO: 55  
; LENGTH: 1510  
; TYPE: PRT  
; ORGANISM: Xanthomonas  
; FEATURE:  
; US-10-418-861B-55

Query Match 4.1%; Score 361; DB 4; Length 1510;  
Best Local Similarity 21.2%; Fred. No. 1.7e-15; Mismatches 304; Conservative 158; MisMatches 466; Indels 504; Gaps 75;  
Matches 304; Mismatches 466; Indels 504; Gaps 75;

QY 144 VLTGEGNNNAFDLK-----VPKKULLNPACHAIYDWNIEATOPRNLNRYDDLG 191  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
250 ILTRENGNFTYTFKKSANGTWTPDMWRETISETQVQDNGTILTGWQVATDSDKEQF--DLDG 307  
Db 192 HDIPELNLNEYQOGLKLKILTLFFPGOKEGYRERLFRLRNQINSH-----F 236  
Db 308 --KUTGISY-----TDGEQILTWTAGRQLQSNTDGRRLFLAYQADRT 350  
QY 237 SLLGNNENPLTWSSFGYTPIGKNGILGOWITSMAPGLKETV--NYSNNNQGHFRPQ--- 290  
QY 351 QVGLPDGMVLAYGDOSARL---RSVTQIOTAGAVASAIIAGDYGN---ARFDALTT 402  
QY 291 -----ANLPVLPVTLMKQVPGAGQPAI-QAEYSY----TSHNYVGG----GSN 330  
Db 403 HRDEQDQVWASWTVQDQQRTRSVHGPCTKIDEATAVSGNTSTVSNALGNAYTRGTS 462  
Db 331 GIWWNKLDNTYGLMTENYNSSESRR---YKDKEATIVR1BERTNNYHLLTSECKQ 385  
Db 463 KLGQAKVTAQVQGLCPCBTVGAFKSKSYDGNGYQDPOEVDFQAVATDWRYNNRGLASKIEA 522  
QY 386 CN--GYIQTETAYVATIGNFNSDQSPOSQFOLPKKTETWRSAD-----NSYRSE 432  
Db 523 DNSGGQKRTLQT-----DWIPS\_FRPV-TDORTYDASDILVARTSWTNSRGQA 570  
QY 433 IETTFDESGLPLTKIKDKTKQK1ISPSTWWEYYPAGWDNCPPPGFFTRVKK1Q 492  
Db 571 LTWSRTPSCGP-TRV-----TTOYRCEDSD-----TAAG-----NC----- 601  
QY 493 TPYDSEBFKDDPEKFLOYRYSITLGSOSHVTUMLKIEHRHSATQOLLNSTLFOQNTDKSBLGRL 552  
Db 602 -----SLPG-----LLASDGAIRDABSTSYV-----YL 627  
QY 553 IKOTECT-----KGENGKTYSVHFKET-YTKDD---THQOSHSHTTDHFPTIH 597  
Db 628 ADDASCTASASTCPHRKGDRKRVNTAIGQTELAYDAGAGRPLSKDNTGIVT--DVTYH 685  
QY 598 -----RSQVTSRVTGRLSPSTDKDVTOMSYDKLGRLLRTINGSTPYANTLYDYL 651

RESULT 9  
US-10-282-122A-76214  
Sequence 76214, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangs  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Carr, Grant

Db 686 PRGWLTASKVKG-----ADASSEBADRTRIYWPNG--LVRQVTOQDGFATFYD-A 737  
QY 652 NNHQDDNRPPFVTTDNGNQLRNEPFGAGRHTVSQCLKDSDGKFYTIHTQOYDEGR 711  
Db 738 HRLTD-----ITDRAHGNTVHYTLNDAGRKVKEITKDAATLK--RTLSRVYNQLG 786  
QY 712 HRP-----YLTNGRQDTPPKVHLMSKSYDNNGOI-----ANTHWS 753  
Db 787 LKTOATAASDPTFPAYDANGNATRKVTDALATATOSEVIMPLNRLSHTLQDVAGTKADTFKA 846  
QY 754 YGVSBKIT-----VDTITLITKQLOSNNSNWQ---TSGEVTVTPSQQPIQITLFBAGH 806  
Db 847 YDALDNLKVTDPKGDLDTYDNGFGBLVKLTSPDGTG--VTSY-----YDSGN 894  
QY 807 LOSCHTLTRDGMDVRKRKETDALGQCTIVQDNYNRVIQITLPGTIVRKYAPPSTDTL 866  
Db 855 RAT-----QDARGENTTAYSYDALNRLKTVTIPS-----SLDVTY 910  
QY 857 T-DIVUNGQISLSGQDPLSRLTOSQDCGRVWATYTSAGNDQCPSTVTPDGOPIHYQO 925  
Db 931 TYDVTQACTSG-BTF-SIGRLKMQDGAITOYCVYCNRGD-LVRKVQTSNGTALVLRD 987  
QY 926 PEL-----DDAVQVASNE--ITQQFSNPVUGALLKAVAEGSOLTPIYPSRL 973  
Db 988 YTVGGQLRMTYFDGAVVTDYVRNAQOGT-QTOGVGTPAGGS--ROYLGLA--TYPFEGA 1042  
QY 974 KMENINDMKKMSIWIIL-----RGLENYUPLTGITQKISDHGRVTOQIDS 1021  
Db 1043 AGWYQGNCRTLARQYDLYRQPQALQPTPQGGLOVGF-----GDPAGNLTALTPA 1092  
Db 1022 SIKIT--LNYDDAHRHGSQVTDATRMLT-----VEFGLNREIGRKUCDSSGHTLQIQ 1078  
Db 1093 GNTPEIGLGYDALGRLTG--LTDGNTGTCV-----DG----- 1123  
QY 1079 SWLKTQQLANRTRVKGNGVQRTQSYNSR-NRINQYKXGDBCPCTDKYGHSTVTONFY 1137  
Db 1124 -----YSYDATGNRL-----STKVG--TATQIVTY 1146  
Db 1138 DIGNITACHTFADGTEDHATFKFANITDPQCOLTEVWHTHPDMDNRLKYGAGRIN 1197  
Db 1147 -----PNSHRLSAV-----AGYARIYDTG--N 1168  
QY 1198 ITDMHGNTRNTVYTLGRQNGQ--GSTYGYDPLNRLVQSQRTDLDCELYRSTMLVHEV 1255  
Db 1169 TTAIGTARQYTDTSGRMTQARRAGAV-----TMARYN 1203  
QY 1256 RNGEMI--RLRTGETTIAQORASKVL-----LTDGDSOSVIL----- 1292  
Db 1204 GRKGEQVRFLGTTNTYTFDEASHWLGYDTSNGAPKQWAIWDDLPVGLANAKHHTB 1263  
QY 1293 -----TSD-----KQMSQEAYASAYGKHKSSTAND--ASILS--YNGERADPVs 1331  
Db 1264 RDHHSQPVVWIDPFRDVAVWTSILKGEAFTGNTAPNQDDPGDGAVALVLMKRPFGORPDAAS 1323  
QY 1332 GVTIHLGNGRSYDITLMLRHTPSLSPSFG--AGGINPVSICLGDPINRSPSG 1382  
Db 1324 GLNQ--NWFDRYEATGRGQSD--PIGEGGISTVSYSSPVKYDVG 1370

APPLICANT: Yamamoto, Robert  
 APPLICANT: Forbyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA-034A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-05  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,309  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SEQ ID NO: 76214  
 LENGTH: 843  
 TYPE: PRT  
 ORGANISM: *Salmonella typhi*  
 US-10-202-122A-76214

Query Match 4.0%; Score 353.5; DB: 4; Length 843;  
 Best Local Similarity 22.8%; Pred. No. 2.3e-15; Mismatches 308; Indels 327; Gaps 48;  
 Matches 223; Conservative 121; MisMatches 308; Indels 327; Gaps 48;

QY 542 YNTDKSELGR---LILQQTCTKGENGKTVSSVHKFTYTRQDPDQLQQSISITHDNFTH 597  
 Db 32 HHTSGEQRDLWMLAERSLVCQVTDSMGRG-----RCHWMDAQLVTR 74  
 QY 598 R\$Q\$RSR\$YRG-----LP\$DPTKDTY\$TOMSYDKLGRLLTRTANS\$GPYMTLYOE 650  
 Db 75 RDEAQGQMT\$FRWSDEERLLIGMTDQGGKWRVYVDRGLHL-----TETHD-P 120  
 QY 651 LNNI\$QDDNRP\$PP--VITTDVNGNQLRNED-----GAGRHVSQCL 689  
 Db 121 LG\$VETQTM\$PWH\$HOPETEVDAA\$GA\$W\$AY\$EY\$D\$B\$G\$T\$H\$Q\$R\$T\$Y\$G\$Y\$R\$H\$-G\$G\$V\$ 179  
 QY 690 K-\$P\$DG\$G\$K\$F\$Y\$T\$H\$Q\$D\$B\$G-----HT\$T\$V\$D\$Y\$T\$N\$G\$R\$O\$T\$D\$P\$K\$H\$MS 738  
 Db 180 R\$T\$D\$A\$G\$G\$K\$Y\$-----Q\$N\$E\$D\$Q\$M\$R\$K\$T\$C\$S\$Q\$T\$A\$W\$Y\$D\$E\$R\$T\$L\$E\$R\$T\$A\$S\$-N\$TR 232  
 QY 739 K\$Y\$D\$W\$G\$O\$L\$A\$N\$T\$H\$S\$Y\$G\$V\$B\$K\$T\$W\$D\$P\$T\$I\$T\$A\$K\$Q\$L\$O\$S\$N\$N\$W\$P\$Q\$G\$K\$E\$V\$T\$T\$P\$S\$Q\$P\$O\$ 798  
 Db 233 Y\$Y\$D\$G\$N\$G\$H\$T\$E\$V\$M\$F\$A\$D\$G\$R\$T\$E\$R\$Y\$Q\$D-----A\$A\$G\$R\$L\$V\$K\$T\$S\$P\$A\$G-----Q\$1 272  
 QY 799 T\$P\$F\$B\$A\$G\$H\$Q\$S\$C\$H\$T\$R\$G\$W\$D\$W\$R\$V\$K\$E\$T\$D\$A\$G\$C\$T\$Y\$Q\$D\$N\$Y\$R\$V\$Q\$T\$U\$P\$D\$G\$T\$V\$N\$R\$K\$A 858  
 Db 273 T\$R\$Q-----R\$D\$G\$G\$R\$V\$R\$O\$T\$D\$A\$T\$G\$R\$T\$A\$Y\$D\$G\$R\$T\$T\$T\$N\$G\$S\$Y\$R\$F\$Y-----Q\$1 319  
 QY 859 P\$F\$S\$T\$D\$T\$L\$T\$D\$R\$V\$N\$G\$S\$L\$G\$Q\$F\$D\$G\$L\$S\$R\$J\$T\$Q\$S\$Q\$D\$-G\$G\$R\$W\$A\$Y\$T\$S\$A\$G\$N\$Q\$C\$----- 909  
 Db 320 -----D\$U\$D\$R\$T\$E\$T\$Q\$P\$G\$S\$R\$A\$Y\$G\$M\$A\$N\$A\$V\$A\$Y\$G\$G\$R\$ 356  
 QY 910 -----S\$T\$V\$T\$P\$D\$G\$O\$P\$H\$Y\$Q\$O\$P\$E\$D\$A\$V\$Q\$-----A\$S\$E\$N\$T\$O\$P\$ 944  
 Db 357 G\$E\$R\$H\$G\$L\$E\$R\$D\$A\$G\$R\$L\$T\$A\$K\$T\$P\$E\$---T\$T\$E\$Y\$R\$Y\$D\$A\$D\$R\$L\$E\$R\$R\$H\$D\$A\$E\$G\$G\$P\$E\$V\$-R\$P 412  
 QY 945 S\$Y\$N\$P\$V\$T\$G\$A\$L\$K\$K\$A\$V\$A\$E\$G\$-----Q\$S\$T\$P\$Y\$P\$G\$R\$L\$M\$E\$N\$D\$K\$M\$S\$Y\$W\$T\$R\$G\$E\$N 995

RESULT 10  
 US-10-202-122A-47600  
 ; Sequence 47600, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Habelbeck, Robert  
 ; APPLICANT: Ohlson, Karl  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forbyth, R.  
 ; APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA-034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-05  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22



Db 218 RQTQFHRERAGEFSGEITGVT-DGAGRHRFLVLTQAORAEARQQATSGGTPEAPSFT 276  
 Qy 462 :-----T-----THWEXTPPAGEVEDNCPEP---RGFT-----RFTV 488  
 Db 277 LPGYTEYGRDNGIRLSAVWLTHDPPEP---ENLPAAPLVRVCGWTPRGELAAYDRST 331  
 Qy 489 KIOTPYSEFK-----DDEPKFIQYRSYLGSOSHVTKIEERHYSATOLLNL 539  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 332 OVSFTYDDKYGRMRMVAHRTGRPE---ICRYD---SDGRVTEQNLNPGALSYT----- 379  
 Qy 540 PQYNTDKSELGRLIKQTEC---TKGENGKTYSVHKFTVKQDDTLQQSHSITHDNFTIH 597  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 380 YQEKRITITDSLNREVLAHTQOEGG-LKRVICK-----EHADGVT 421  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 598 RSOVRSRYSYGRLESDUTDKDITVOMSYDKLGRLTRTLNSGTPANTLYDVEILNLQDD 657  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 422 OSQFDA---VGRLRAQTDAGRTTEYSPDVWTLGLTRIT---TPDGRASAFYY----- 468  
 Qy 658 NRPPFVITTDVNGNOLRNFBDGAGRHSQCLKSD---GDCKPYTHTQQDEQRH 712  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 469 NHHSQSLTSATGPGLUEIRBYDENGRLQBTAPGDITRYRVNPNSDLPCTEDATGSR 528  
 Qy 713 HTSITSDYLJNQROTDPKVHLMSMSKSYDNNGOTANTHWSGYSEKLTVDPL-TLTWK 771  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 529 KTMWSRY--GQLSFTPCGSYTRYDHDFRGQVTAHREGISQRAYDSRQQLIAK 585  
 Qy 772 QLOSNSNNYTOTGKEMVTTYRSPQBLIOITLPDEACHLQSCHTLTDGWRVRKETDAQO- 830  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 586 DTCQHETRYE-----YNAAGDLTT---VIAPDG-SRNGTQYDAGKA 623  
 Qy 831 -CII-----YQDNMYRVIQITLPDGITVNRKXAPSTIDLTIDIRUNGISLGGQTP- 881  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 624 ICTTQGLLRTMSMEYDAASGRVILRPSENG-----HTTERRYDVIDRLQBTGFGRTORH 678  
 Qy 882 -DGLSLRITOSQDGRVWAATYSAGNDQCPISTVITPDQFIHYQYQPELDDAVLQVASNEI 940  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 679 HDTGKLLRSEDEGIY-----THWHD-EADRHTHTRVGET 714  
 Qy 941 TQOFSYNPVGAL--LKAVAEQSILPTY-YPS-GRLKMENTI-----NDM---KKSY 986  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 715 AERMQYDE-RGWLTDISHISEGHHRVTVHGYDSKGRLASEHLTVHHPOTNELLWQHETRH 773  
 Qy 987 LWTIIGLENGYTDTGCTQKISRTHGRVQIKOSSIKITLNDLNHIGSQVDTATG 1046  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 774 AYNAQGLAN---RCIPSLPAVENLYG-----SGMUSGMKLG 808  
 Qy 1047 HMLTTTVEE-----DGIANREIGERKLCDSSGGHTUDIOSWLUKTQQLANRIVKLNGVQRTQEOY 1104  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 809 D-TPLVETTRDRHLRETIR---SFGRYELTRATTAGQLOSQ---HNSLIS-DRYD 858  
 Qy 1105 YDSRNRLNQYKDCGACPTDKYKGHSIVTONCPFTYDIGNITACHTFAD---GTEDHA 1158  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 859 WNDNGELIR---ISSPRQ-----TRSYSYTTGRLLGVTAAIDLIRIPTYTDPA 906  
 Qy 1159 TKPKEANPTDPCQLTEVHHTHPD---MDN-----IRLKQDAGRVINTD----- 1200  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 907 GNRLPDP-----ELRUDSALSWMWDNRDARTADAHLYRDRHGRLTKEUDLIPEGV 956  
 Qy 1201 --NHGNTENTFXTLGLRQLONGGSVYG-----YDPLNRLVSQKDTDCEL--- 1244  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 957 IRTDDETERHYHVSQHRIVHYTQYAABLPLVRSVLDPGRVAKRVRERDLCW 1016  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Qy 1245 -----YY-----RETMVNE-----VRGEMIRLLR---TCBTIIQORA--- 1276  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 1017 SLSRKPKQVTWGMDGRLLTQNDRTRICITYQPGSFTPLRIVETATGBLATORRSLAD 1076  
 Qy 1277 -----SKVLTGDSQSQSMTLTSKQNLQEA----- 1303  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 1077 TLOOGSGEDOGGSUVFPPLVQMLDRLESILA---DRVESSRMLASCGLTVAOMOSM 1133  
 Qy 1304 -----YSAVGKHKSTANDAS-----ILGNG 1324  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 1134 DPVVTTPARKIHLVHCDHRLGULPLALISTECTTAWYAEYDEMGUNLNEENPHQLOQQLIRPG 1193

Qy 1325 ERADPVSGVTHLNGYRSYDPTLMRFHPPDSISPPFG-AGGINPYSYCIGDPINRSDESG 1382  
 :-----|:-----|:-----|:-----|:-----|:-----|:-----|:-----|:  
 Db 1194 QDIBEEGLEY-NRHRVYDPLQGRYITQD---PIGLKGWNFQYPL-NPISNIDPLG 1246

RESULT 12

US-10-282-122A-43060

; Sequence 43060, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Lianguo

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlson, Karl

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELTRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

PRIOR FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 601206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 601207,727

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 601253,335

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 601257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 601242,578

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 601267,636

PRIOR FILING DATE: 2001-02-16

PRIOR APPLICATION NUMBER: 601257,931

PRIOR FILING DATE: 2001-02-16

PRIOR APPLICATION NUMBER: 601267,636

RESULT 13  
US-09-912-020-340  
; Sequence 340, Application US/09912020  
; Patent No. US2002045592A1  
; GENERAL INFORMATION:  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Trawick, John  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Froelich, Jamie M.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
; TITLE OF INVENTION: ESCHERICHIA COLI  
; FILE REFERENCE: ELITRA.001DVL  
; CURRENT APPLICATION NUMBER: US/09/912,020  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: 09/92,709  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: 60/117,405  
; PRIOR FILING DATE: 1999-01-27  
; NUMBER OF SEQ ID NOS: 485  
; SOFTWARE: FastSBQ for Windows Version 3.0  
; SEQ ID NO 340  
; LENGTH: 1426  
; TYPE: PRT  
; ORGANISM: *E. coli*  
US-09-912-020-340



; PRIOR FILING DATE: 1999-01-27  
 ; NUMBER OF SEQ ID: 485  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO: 340  
 ; LENGTH: 1426  
 ; TYPE: PRT  
 ; ORGANISM: E. Coli  
 ; US-10-771-241-340

```

Query Match 3.8%; Score 339; DB 5; Length 1426;
Best Local Similarity 20.2%; Fred. No. 5.5e-14; Gaps 60;
Matches 207; Conservative 150; Mismatches 409; Indels 458; Gaps 60;
Match 1090 ---IVKLNGVLTQRETYQSISRSRNLRNQYKCDGAEC---PTDKYGHST----- 1130
Match 812 GGTGPVLYETVRDRHRETVRSPGSMAGSNAYELTSYTYPAGOLQSOHLNSLVYDRYGWS 871
Match 1090 ---IVKLNGVLTQRETYQSISRSRNLRNQYKCDGAEC---PTDKYGHST----- 1130
Match 872 DNGDLVRISGRQ-TREYGYSATGRLSVRTLAPDLDIPIYATPDAGNRLPDPBLHPDS 930
Match 1131 -----VTON---FTYDIGNITACHTTPADG---TEDHATFKFANPTDCQCQLEV 1174
Match 931 TLTWVDPNRRAEDAHVYRVHDYGRLEKTRKDPIAVGVTDBERT----- 975
Match 1175 HHTHDMPNDNIRLKDKAGRVINID-NHGN---TENFTVTLGR----- 1215
Match 976 HHYH-----YDOSHLRVYFRIOKGHEPVLVSSRYLDPGLGRMAKRWRERDTG 1025
Match 1216 ---LQNGQSVYGDPLNRLIVSQKUDTLCDELYR----- 1247
Match 1026 WMSLSRKPEVTWYGMG-DRLTVQDITRIQTVPFGSFTPLIRVETENGREKAQRSS 1044
Match 1248 --ETMLVNEVTRG-----EMIRIL-RTEBTIA----- 1279
Match 1085 LAETLQEGSBENGHCJ3VFPABLVRLDRLETRADRVSSESRAWLAQCSLTVEQLAROV 1144
Match 1280 LLTGDSQS-----VILUSDQKQISQEA-YSAVGKHKSTANDASILGN--- 1323
Match 1145 EPEYTPARKAHLYHCDHGRGLPLAL-SEDGNTAWSAYDEWGNQNLBENPHV-YQPYRL 1202
Match 1324 -GERADPVSGVTHLQLENGYRSYDPTLMRFPHDSDLSPFG-AGGINVSYCYLGDPINRSDPS 1381
Match 1203 PGQQDDEESGUYY--NRHRYVDPLGORYTQD--PMGLKGQWNYQYPL-NPQIDPM 1226
Match 1382 GHLSHQAWTGIGMGIAGLLTITATGMAIAAGGIAIAIASTTALAFGLSVSDITS 1441
Match 1257 GLL-OTWDDARSG-----ACTGSV---CGVJSRIGPSKFESTADALD----- 1295
Match 1442 IVSGALEDASPDKASISLIGWYMSMGMAAGLBSAIIKGKTIAHLSAFAEDEGENALKSTS 1501
Match 1297 ---ALKETONRS-----LCDMNEVSGIVCKDTNG-----KYFASKAETDNLR-K 1336
Match 1502 ESSRKWGVNTRSLDRBIVRNREGQVIKDHSGRGYTNFMGKGEQAIIVHGK--DPLYHT 1559
Match 1337 ESYPLRKRCPGTDRVAYHTG--ADSHGHDYVDFPSSSDKNLVRSKONNLEAFYLAT 1393
Match 1560 EGNKH---NGKGPY 1570
Match 1394 PDGRFEALNNKEY 1407

```

**RESULT 15**  
 US-10-771-241-340  
 ; Sequence 340, Application US/10771241  
 ; GENERAL INFORMATION:  
 ; Publication No. US20040241715A1  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Forsyth, R. Allym  
 ; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
 ; TITLE OF INVENTION: ESCHERICHIA COLI  
 ; FILE REFERENCE: ELTRA\_001C1  
 ; CURRENT APPLICATION NUMBER: US/10/771, 241  
 ; CURRENT FILING DATE: 2004-02-03  
 ; PRIORITY APPLICATION NUMBER: 09/492, 709  
 ; PRIORITY FILING DATE: 2000-01-27  
 ; PRIORITY APPLICATION NUMBER: 60/117, 405



**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
 Copyright (c) 1993 - 2006 Compugen Ltd.

On protein - protein search, using SW model  
 Run on: January 30, 2006, 09:46:19 ; Search time 13.3373 seconds  
 Scoring table: BLOSUM62

Title: US-09-889-874A-23  
 Perfect score: 8879  
 Sequence: 1 VYIKPLKLFRRITMSDNNEP.....PRKILIGRTEKTVKPKTFRP 1673

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Published Applications AA\_New:  
 1: /cgn2\_6/ptodata/1/pubpaas/U508\_NEW\_PUB.pep:  
 2: /cgn2\_6/ptodata/1/pubpaas/U506\_NEW\_PUB.pep:  
 3: /cgn2\_6/ptodata/1/pubpaas/U507\_NEW\_PUB.pep:  
 4: /cgn2\_6/ptodata/1/pubpaas/EFCI\_NEW\_PUB.pep:  
 5: /cgn2\_6/ptodata/1/pubpaas/U509\_NEW\_PUB.pep:  
 6: /cgn2\_6/ptodata/1/pubpaas/U510\_NEW\_PUB.pep:  
 7: /cgn2\_6/ptodata/1/pubpaas/U511\_NEW\_PUB.pep:  
 8: /cgn2\_6/ptodata/1/pubpaas/U560\_NEW\_PUB.pep:  
 9:

Pred. No. 19 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	285	3.2	2721	7 US-11-096-051-10
2	285	3.2	2725	7 US-11-096-051-8
3	284.5	3.2	2715	7 US-11-113-424-51
4	283	3.2	2376	7 US-11-096-051-4
5	283	3.2	2715	7 US-11-096-051-2
6	273.5	3.1	1094	6 US-10-821-234-1097
7	272.5	3.1	2333	6 US-10-453-372-170
8	272.5	3.1	2662	6 US-10-453-372-114
9	272.5	3.1	2724	6 US-10-453-372-148
10	272.5	3.1	2733	6 US-10-533-372-136
11	272.5	3.1	2733	6 US-10-453-372-142
12	272.5	3.1	2733	6 US-10-453-372-146
13	272.5	3.1	2733	6 US-10-453-372-150
14	272.5	3.1	2733	6 US-10-453-372-1154
15	272.5	3.1	2765	6 US-10-453-372-1148
16	271	3.1	259	6 US-10-453-372-168
17	265.5	3.0	2515	7 US-11-113-424-53
18	253	2.9	1045	7 US-11-113-424-54
19	253	2.8	2669	7 US-11-113-424-14
20	252	2.5	2725	7 US-11-113-424-52
21	216	2.4	2053	7 US-11-013-759-9
22	203.5	2.3	2314	7 US-11-013-759-11
23	195.5	2.2	1579	7 US-11-052-554-9
24	194.5	2.2	5024	6 US-11-052-554-91
25	2902	7	US-11-052-554-91	

RESULT	1
US-11-096-051-10	; Sequence 10, Application US/11096051
	; Publication No. US20050244868A1
	; GENERAL INFORMATION:
	; APPLICANT: Kerkuta, Ramesh
	; APPLICANT: MacLachlan, Timothy K
	; APPLICANT: Rastelli, Luca
	; APPLICANT: Vernet, Corine
	; APPLICANT: Ettemberg, Seth
	; TITLE OF INVENTION: Tet-M3 Polypeptides and Polynucleotides and their Methods of Use
	; FILE REFERENCE: Attorney Docket No. Ciria 967
	; CURRENT APPLICATION NUMBER: US/11-096,051
	; PRIORITY APPLICATION NUMBER: 10/038,854
	; PRIORITY FILING DATE: 2001-12-31
	; PRIORITY APPLICATION NUMBER: 60/557,978
	; NUMBER OF SEQ ID NOS: 38
	; SOFTWARE: CuraseqList version 0.1
	; SEQ ID NO: 10
	; LENGTH: 2721
	; TYPE: PRT
	; ORGANISM: Homo sapiens
	US-11-096-051-10
	Query Match Best Local Similarity 3.2%; Score 285; DB 7; Length 2721; Matches 367; Conservative 253; Mismatches 637; Indels 626; Gaps 97;
QY	18 NEFTQANIFTSAVGGGWDPTQLNQIQTGLHIVNGNGLGTLPLTISYSFLNKIDGP 77
DQ	1151 NORISDOPPWVSIINGRSRISCPSCNQADGSKLLARVA-----LAC 1195
QY	78 GI-----GEFGLSLVYDRKN--SLSLS-----TGENYKVE 107
DQ	1196 GIDGSLSVGVDFNVRIRFFPSGNVTSHLRKDQFRISSNPARYLATDPYGDY--VS 1253
QY	108 TDCKVKGQOKKUDNLRFKDLGNCYCILHKSGDIEITLGEMNA--FD-----LKV 157
DQ	1254 DTWTRRKYRK--SLTGAKDLTKNA-----EVAGTGBCQCLPDEARCDDGKAV 1301
QY	158 PKKLNQPGHAAYIDMFEATOPRLNLYDGDGHOPILPLNLYQGLKTILTPQKE 217
DQ	1302 EATIMSPGMAY-----DKQNLIVFVDG-TMIRKVDONGISTILL--GSDN 1344

Sequence 283, App  
 Sequence 90, App  
 Sequence 92, App  
 Sequence 3, App  
 Sequence 13, App  
 Sequence 4, App  
 Sequence 5, App  
 Sequence 7, App  
 Sequence 281, App  
 Sequence 3188, App  
 Sequence 2, App  
 Sequence 202, App  
 Sequence 50, App  
 Sequence 2, App  
 Sequence 171, App  
 Sequence 6, App  
 Sequence 3154, App  
 Sequence 6, App  
 Sequence 200, App  
 Sequence 131, App  
 Sequence 248, App

QY 218 GYRTTELRFLNQRQLNSIHNFLSGNENPLTWSPGTPIGKNGILGOWISMTAPGLKSTVN 277  
Db 1345 -----LTSARPLCD--TSMHISQREWPTL-AINPMDNISIY 1380  
QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 1381 VLDNNVVLQITENQVRIAAGRPMHQFGVEYPVGKHWQT--TLESATAVSYGV- 1437  
QY 334 NUNKDNLYGLMTEYNGSTESSRMVKDKEGHDDQIVRIERTYNNTHLTSRCKQN----- 387  
Db 1438 -----LY--ITERDBKKINRIRQVTDTGEISLVA-----GIPSBCDCRNDAACDC 1480  
QY 388 -----GYIQTE-----TAYAIGH-----NF-----D 406  
Db 1481 YQSDGIVYAKDAAKUSLAAPSSLAASPGTLYTADLGMRRAVSKNPKLNSMNFEVASPTD 1540  
QY 407 SQPSQFOLPKTKTETRSADNSY-----RSEBTTETFDSENPLTKVKKD-KKTQK 456  
Db 1541 QELYIFDINGTHTQTVSLVTDYLYNFSISNDNDITAVT-DSNANTL-RIRPNRMVR 1598  
QY 457 IISPSHT--WEVYPAGEVDNCPPBPGFTRFVKKIIQRPYDSEBFKDPPEKFIQYRYSIL 514  
Db 1599 VVSPDNQVWLTIGTINGCLK-----GMTAQGLBLVLFYH----- 1633  
QY 515 GSOSHUTLKIERHYSATOLNISTFLQMTDKSELGRULKQTC-----KGNGKTY 588  
Db 1634 GNNSGILATKESDETGW-----TTFDVSDE---GRLTNTFPQGVNLHGDMKLT 1682  
QY 569 VVHKFETYTKODDTQOOSHSTITHDN-THRSQVRSRYTGRLFSDTDTKDIVTOMSYDKL 627  
Db 1683 VDIESSREEDSVI--TSWISIISFSFTWQDQQLANSY-----QIGYGS 1725  
QY 628 GRLLTRTLNSGTPYANTLT--YDVLLNNOQDNRPPFV--ITTDVNGNOLNEFDGAG 682  
Db 1726 LRII-----VASGLDISHQYEPHVLAGTANPTVAKNMILPGENGONL----- 1768  
QY 683 RVHSOCOLKOSDGCGFYTHTQDQEGRHTHTSDYLJNGRQOTDPKWHMSKSD 742  
Db 1769 -----VWRFRKQAOQK--UNVFRKLURGNMLSVDFDTTKEKLYDDHKKFL-LRIAYD 1824  
QY 743 NWGQIANTHNSYGVSEKITIDPITATKOLQSNNSNNAVQTGKEVITYPSQOPIQITLED 802  
Db 1825 TSGH--PTWLM--PSKLMANUVNTVSSTGQAS-----IDRG-----TWSK----VVD 1866  
QY 803 EAGHLSQSCHTTRDG--WDRRKENTAI----GOCTIVQYDNKNRVIQITLPLD----- 849  
Db 1867 GQGRIVS--RVPADGKWTWSITYLEKSMWLLHSQRYIFFDMDRISAITMPVARHIM 1924  
QY 850 -----GTIVNRYKVAFFPSTTDLITIRVNGISLGQQTDCISRLTQSQDGGRWAYTSA 903  
Db 1925 QTIRSTGYVRNLYNPESNASIITDYNEGILL-OТАFLGTSR-----RV----- 1968  
QY 904 GNDQCPSTVITPDQPHIYQYQPEELDAVIVQASNBITQOFSYNTVGALLKAV--AEG 960  
Db 1969 -----LFKYRQTRLS---EILYDPSTRYSTFYTDETAG-VLKTVNLQSDG 2008  
QY 961 OSLTPTYPPSRGLKMENINDKMSLWLTGLENYTDLGTGIOKISRDTGHRVTOIKD 1020  
Db 2009 FICTIRYRQIGPL----IDRQFRPSELGMVNARFD----YSDVNSPRVTSMQG 2054  
QY 1021 SISKTL-----NYDNLNRHGS-----QVTDLATHMLTTVEFD--GLAREIGR 1064  
Db 2055 VINETPLPIDYQFDDISGKVQFGKRGVIVYDINOIISTAVMTKHFDAHRIKEIQ 2114  
QY 1065 KLCDSGHTLIDQSQSLKTOQLANRVIKNGVLUQREFOQYPSDRSRUQLAQKCDGAECPT 1124  
Db 2115 EIFRSIAMYWTLQYD--NMGRVTKRISKI-CPPANTKTYA-----BYDVGD-OLQT- 2162  
QY 1125 KGHSIYTQNTFYDIVGNITACTTFADGTDHATPKFANPTDQOLTEVHHTHPDMPDN 1184  
Db 2163 VYLNKEIMRRTYDNLN-----HLLPSNSARLTP----- 2194  
1185 IRLKYDKAGRVINITD-----NHGNTENFTYDTLGLRQ-----NGQSVVYGYD 1227

Db 2195 --LRYDIDRTRIGDQVYRQRLDEDGLRQRTFEVSKGLITRVSKGSGWTVIYRD 2252  
QY 1228 PLNRLVSOQT-----DTLDCLYY--RETMVUNEVNGE 1259  
Db 2253 GLGRVVEKSLSQGHLQFPPYADITYPTRTRVYHNSSESRTSILYYDQGLHAMEISGGD 2312  
QY 1260 MIRLRTOBTITAQRAKSVLJGTDQSOSVILTSDKONLSQEASVAGK--HKSTANDA 1317  
Db 2313 -----EFYIASDN-----TGT-----PLAVFSSNLMLKQIQTAYGEIVFDSNIDFQ 2355  
QY 1318 SILGYNGRADPVGVTILGNGYRSYDPTLMRFTPD-----SISPG----- 1360  
Db 2356 LVIGFHGGLYDPLTKLHFGE--RDYDILAGRWTPEIWKRIGKOPAPFNLYMPRNN 2413  
QY 1351 -----ARGINPSYCQDPIPNSD--PSGH--SWQWTGIGI 1396  
Db 2414 PASKHDYDYLTDVNLSVLUVTGFLHNAIPGPVVKDLETERSYELVSKQOQDDI-BPI 2472  
QY 1397 AGULLATGGMIAAGGIAALAS-----TSTIALRGALSVTSD--I 1439  
Db 2473 FGVQOQVVRQAKFLSLGKMAEVQSVRAGAQSWLWFATVSKLIGKVMlavSQRV 2532  
QY 1440 TSTVSGALDASPKASSTIG-----WWSMGMGAGLASSAIKGTK 1480  
Db 2533 TNVIANIANEDCI-KVAALVNNAPYLENHLIFTEGKDTYFIKITPESDGLTURTSRK 2591  
QY 1481 LATLIGAFAEDGENALLISTSSRSRRIKGVYR-----SLDREVR 1520  
Db 2592 AL-----ENGIVNTVUSSTTUVNGRTRKTFADYEMQFGALAHRYGWTLDERK 2641  
QY 1521 NEEQVQIKHRSRGYTDNMKGKQRAILVHGKDGFLYTEGENHN---GK---GPYT 1571  
Db 2642 ILBQARQARAWA-----REHQVRDUDGEEGARLW-TEGEKROLISAGKVOQGDGYV 2694  
QY 1572 RHTPE--QDVYKDNNTVLDTQ 1592  
Db 2695 LSVEQYPLAD--SANNQFLRQ 2715

RESULT 2  
US-11-036-051-8  
Sequence 8, Application US/11096051  
Publication No. US20050244868A1  
GENERAL INFORMATION:  
APPLICANT: Kekuda, Ramesh  
APPLICANT: MacLachlan, Timothy K  
APPLICANT: Rastelli, Luca  
APPLICANT: Vernet, Corine  
APPLICANT: Ettenberg, Seth  
TITLE OF INVENTION: Ten-M3 Polypeptides and Polynucleotides and their Methods of Use  
FILE REFERENCE: Attorney Docket No. Cura 967  
CURRENT APPLICATION NUMBER: US/11/096,051  
CURRENT FILING DATE: 2005-03-30  
PRIOR APPLICATION NUMBER: 10/038, 854  
PRIOR FILING DATE: 2001-12-31  
PRIOR APPLICATION NUMBER: 10/455, 772  
PRIOR FILING DATE: 2003-06-04  
PRIOR APPLICATION NUMBER: 60/557, 978  
PRIOR FILING DATE: 2004-03-30  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: CuraseqList version 0.1  
SEQ ID NO 8  
LENGTH: 2725  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-11-096-051-8

Query Match 3.2%; Score 285, DB 7; Length 2725;  
Best Local Similarity 19.5%; Pred. No. 4.49-10;  
Matches 367; Conservative 253; Mismatches 637; Index 626; Gaps 97;

Db 1155 NQPISSQQPVVUSSING--NGRRRISICSPCGNGQDGNKLLA PVA-----LAC 1199  
 QY 78 GI-----GPNIGLSVDRKN--SLLS-----TGENKYKE 107  
 Db 1200 GINGSLVYGDFNTYRRIPESGNVTSLERLNKOPRHSSNPARHYYLTADPVGCDLY--VS 1257  
 QY 108 TDKVKGLOOKEDLNUFEDKENCYRITHKSGDIEVUTGFFNNA--FD-----LKV 157  
 Db 1258 DTNTRRTRPK--SLTGAKDLTKA-----EVVAGTGEOLCPFPBARCGDGKAV 1305  
 QY 158 PKKLNPAGHAIYDWNFRATQPPLNRTYDLDCHDIPLNLRQGLKTTLFPSOKE 217  
 Db 1305 EATLMSPKMVA-----DRKGLYTFVPG--TMKRDQDNGISSL-----GSND 1348  
 QY 218 GYTERLFLARLQLNSHNFSNLGENPLWTFSPGTPIGKNGILGWOMTSMTAPGLKEVN 277  
 Db 1349 -----LTSARPLCDD---TSMHISQVREWPML-ANPMDMSI 1384  
 QY 334 NKNUKNLYGLMTEYNGESTESRYKDKECHDQIRIERTYNTNHLSRKQDN-----387  
 1442 -----LY-ITERDEBKKNRIRROVTTDGEISLVA-----GIRSBDCDKDANCDC 1484  
 QY 388 -----GYOTTE-----TAYAIGH-----NE-----D 406  
 Db 1485 YOSGDGYAKDAKUSAPSSLAASPGTLYTADLGMRIRAVSKONPLASMNFEVASPD 1544  
 QY 407 SQPSQFOLPKTKTETWRSADNSY-----RSSETTEFDENPLTKVIKD--KKTOK 456  
 Db 1545 QELYIFDINGTHQTVSLVTDGLYLYNFYSNNDITAVT-DSNONTL-RIRDPNRMVR 1602  
 QY 457 IESPSTH--WEYEPAGEVDNCPPPYGTRFVKILICPYDBRFKOPKEFOYRSLI 514  
 Db 1603 VVSBDNQVTLWLTICGCLK-----GMTAQGLELVLFTYH-----1637  
 QY 515 GSOSHVTLKERHYSATOLINSLFQNTDKSLLGRLUKQECT----KGENGKTY 568  
 Db 1638 GNSGLATATESDECTW-----TFFDUSE---GRUNNTNPFVWNLHGMDKAIT 1686  
 QY 569 VVHKFTYTKQDQHQSHTTNDNF-THRSQVRSRYTGRLFSDTDKDIVTOMSYXKL 627  
 Db 1687 VDLESSRREBDSI-TSNSLSDSFYTVMQDQLRNSY-----QIGDGS 1729  
 QY 628 GRILTRTJMSGTPTVANTL--YDEELNNLQDDDRPFPY--ITTFDUNGQNOLNEEDAG 682  
 Db 1730 LRIT-----YASGLDHSYQTEPHVFLAGTANTPAKRNMTLPGENGONT-----1772  
 QY 683 RHYSQCLKSDGDKFYHTHQVDEQRGHHTSYSDYLNGRQQTDDKHMSMSKED 742  
 Db 1773 --VWRPRKEAQOCK--WVFGKURVNGNLNLSDFDRKTKEYDHRKE-LRTAYD 1828  
 QY 743 NWGQIANTWMSYGVESEKITVDPITUTATKOLQSNNSNNVOTGKEVTTYPSQOPTQITLFD 802  
 Db 1829 TSGH--PTML--PSSKLMANVNTYSSICQIAS-----IORG-----TISEK--WDYD 1870  
 QY 803 EAGHQSCLTTRIG--WWRVKETDAI----GOCTIQYDNTRNVIQITLFD-----849  
 Db 1871 GQGRIVS--RVFADGKTWSYTYLERSMWLILHSOROFYEYDMDWDRLSAITMPSSVARHM 1928  
 QY 850 -----GIVNRKTAFPSIDLTIDIRVNGISLQGQTFQGSLRUTQSODGGRWAWTSA 903  
 Db 1929 OTIRSIGYRNKVNPPESASLTDYNEBGLL-QTAEGTST-----RV-----1972  
 QY 904 GNDQCPSTVITPDQFIFIHQYQPELDAVILQVASNEITQFSYMPVUTGALLKAV--AEG 960  
 Db 1973 -----LFCYRQRQTRLS---EILYDSTRSPTBETAG-VLKTNLQSDG 2012  
 QY 961 QSLTPIYPSGRKMENTINDMKKNSYLTIRGLNGYDILTGTQKISRDRHTERVTOKD 1020

Db 2013 FICTIRTQIGPL-----IDRQIFRRESEDGMNARD-----YSYDMSFRVTSMQ 2058  
 QY 1021 SSTKTT-----NYDDPLRHIGS-----QVTDATAGHMLTTVFEF--GLNREIGR 1064  
 Db 2059 VINBTPLPBDYQFDDSGKVEOFKGKFWVYIDINQISTAVNTYTKHFDAGRIKELQY 2118  
 QY 1065 KLCDSGGTLDIQQSWLKTQQLANRIVKUNGVLQRTQYSTDNRNRLQYKCDGAECPTD 1124  
 Db 2119 EIRSLMMWITQYD--NMGRVTKRETKI-GFANTUTKAY-----EYDVG-QIQT- 2166  
 Db 1125 KYHSIVTQNFTYDIGNITACTTFADGTEDHATFKFANPDTPCQJTEVHHTHPDMDN 1184  
 QY 2167 VLNKIMKMRINYNDL-----HLNPNNSARLTP-----2198  
 Db 1135 IRKXDKAGRVRVINITD-----NHGNTENFTYTDLGRQ-----NGCSGVYQGD 1227  
 QY 2257 GLRRVRSKTSLSQHQLQFYADLTYPTTRITHVNHSSETSLSVYDQHFLAMELSGGD 2316  
 Db 1250 MIRLRTCTTIAQORASKVLLTGTDSQSVILTSKDNLSOBAYSAVGK--HKSTANDA 1317  
 QY 2317 -----EFYIASDN-----TGT-----PLAVFSSNGMLKQIQTAVGLEYFDSNIDFQ 2359  
 QY 1318 SILGYNGGRADPVSGLTQGNGYRSVPPTLMFHTP-----SISPEG-----1360  
 Db 2360 LVIGFHGKXQYDPLTKLKHGE--RDYDILAGRWTPOLEIWKRIGKOPAPENLYMFRRNN 2417  
 QY 1361 -----AGGINPSYCYGLDPINRSD--PSGHU-SWAQWTGIGKI 1396  
 Db 2418 PASKHDYKDITYDUNSNLVTGFHLENPAIRSPVPKDLTSPSYELVKSQWDDI-PP1 2476  
 QY 1397 AGULITITGGMATAAGGIAALAS-----TSTTALAFGALSVTSD--I 1439  
 Db 2477 FGVOQQVARQAKAFSLGKMAEVQVSRRAGGAQSMWLPATVSLIGKGMALAVSQGRVQ 2536  
 QY 1440 TSTVSGAEDADSPKASSITG-----WWSMGAGLAGBSAIKGTM 1480  
 Db 2537 TNVNLNIAHEDCI-KVAAVLNNAFYLENHIFTEGKDTYFIKTTPESDLGTLRLTSCK 2595  
 QY 1481 LATHIGAAEDGENALXKTSSESSRIKGVTR-----SLDREVR 1520  
 Db 2596 AL-----ENGINTVSPSTTUVGRTRFADYEMQFGALAHVRYGMIHLDEBKR 2645  
 QY 1521 NEBGQVIDKDHSGRTDNFMGKGQOALVHGDKDGFLYMEGENKN--GK-----GPT 1571  
 Db 2646 ILEQARQDALARAWA-----REQRQRTDGEERARW-TEGERKQQLISAGKVQGYDGYV 2698  
 QY 1572 RHTPB--QLVYDVKDNNTVLDQ 1592  
 Db 2699 LSVEQYPLBLAD--SANNIQFLRQ 2719

RESULT 3  
 US-11-113-424-51  
 ; Sequence 51, Application US/11113424  
 ; Publication No. US20050260713A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gangoli et al.  
 ; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 2102-225  
 ; CURRENT APPLICATION NUMBER: US/11/113, 424  
 ; CURRENT FILING DATE: 2005-04-21  
 ; PRIOR APPLICATION NUMBER: 60/256, 704  
 ; PRIOR FILING DATE: 2000-12-19  
 ; PRIOR APPLICATION NUMBER: 60/311, 590  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: 60/257, 314  
 ; PRIOR FILING DATE: 2000-12-20  
 ; PRIOR APPLICATION NUMBER: 60/311, 613  
 ; PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: 60/315,617  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/307,506  
; PRIOR FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 60/322,358  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/294,075  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: 60/288,153  
; PRIOR FILING DATE: 2001-05-02  
; NUMBER OF SEQ ID NOS: 190  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 51  
; LENGTH: 2715  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-11-113-424-51

Query Match 3.2%; Score 284.5; DB 7; Length 2715;  
Best Local Similarity 19.6%; Pred. No. 4.7e-10;  
Matches 302; Conservative 214; Mismatches 524; Indels 497; Gaps 80;

QY 18 NEFFPQANNFTSAVGGVDPRTGLNQIQTGHIVGNGLGPTLPLTYSPLNKTDIGP 77  
Db 1145 NQF1SQQQPPVSSIMG--NRRRS1SCPSNGQADGNKULAPVA-----LAC 1189  
Qy 78 GI-----GRFNGLSVYDREN--SILSLS-----TGENKYTE 107  
Db 1190 GIDGSLYVGDFNYVRKIFPGSGNVTSVLELRNDFRHSSNPARYTLATDPVTGDL--VS 1247  
Qy 108 TDKWKLQKETLDNRKEKENCYRILHSGDIEVUTGFFNNNA--FD-----LKV 157  
Db 1248 DTNTTERYRPK--SLTGAKDTKA-----EVAGTGEOPCLPDEARCDDGKAV 1295  
Qy 158 PFKKLNPAGHAIYIDWNFEATQPLRNRTDDGDHIDPLNLQEVQLIKLTLPFGKE 217  
Db 1295 EATLMSPKGMAI-----DINGLYFVDG--TMRKVDQNGLILSTL---GSND 1338  
Qy 2118 GYRTTELFLNRLNQNSIHNSLGNENPLTWTSFGYTPIGKONGILGOMITSMTAPGGLKETVN 277  
Db 1339 -----LTSARPLTCO--TSMHISQVRLEMPTDL-AINPMDSIV 1374  
Qy 278 YSNNNQGHFPEFQSAMLPVLPYVUTLNUKQVCG---AGOPAQAEVSYTSINYVGGSGNIGW 333  
Db 1375 VLDNNVVLQQTENRQVRIAAGRPMHCQVPGVEVPVKHAVQT--TLESATAIASVSYGV- 1431  
Qy 334 NNIKLDNLYGIMTEYNGSTERRYKODKEGHDQIVIERTVNNHILSTCKQON-----387  
Db 1432 -----LY--TTEDBKKNIRIROVYTDSBISLVA-----GIPSBCCDKNDANCDC 1474  
Qy 388 -----GYIQTTE-----TAYVIGH-----NF-----D 406  
Db 1475 YQSGBOYAKOKAQNAPSSLAASPDGTIXLXIALGNIRARSKNKLNSANFVEVASPD 1534  
Qy 407 SOPSOPOLPKTKTETRSADNSY-----RSBTETTEDESGMPPLKTYKD--KKTOK 456  
Db 1535 QBLYIFDINGTHQYTVLSVTGDLYNFSYNDNDVIAVT-DSNGSTL-RIRDPNMPVR 1592  
Qy 457 1LSPSTH--WVYPPAGEVNDCCPBPYGFTRFKKLIQTYDSEKDD-BEKFQYRSLI 514  
Db 1593 TVSPDNQVIMLTGTINGCLKSMTAQL-----ELVIFTYH-- 1627  
Qy 515 GSOSHVTLKBERHYSATOLINSLRQYNTDKSELGRLLKQTECT-----KGENGKTS 568  
Db 1628 GNSGLLATKSOTEGW-----TRPDYDBS--GRITWTPEPGVUMLNGMDKAIT 1676  
Qy 569 VVHKFHYTKDQDLSHSTTHDNF-TIRHSQVSRYTGRFLPSDMDTKDQVTONMSYKL 627  
Db 1677 VDISSREELSVI--TSMLSIDSFTWMODLRNSY-----QIGYDS 1719  
Qy 628 GRLLRTLNSCTPYANTLT--YDVEUNLQDNRNPFV--ITTDVNGNQLRNEFDGAG 682  
Db 1720 IRIF-----YASGLDSHVOTEPHVLACTANPTVAKRINTLPGENGONL-----1762

RESULT 4  
US-11-096-051-4  
; Sequence 4, Application US/11096051  
; Publication No. US20050244868A1  
; GENERAL INFORMATION:  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Macchian, Timothy K  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Vernet, Corine  
; APPLICANT: Ettenberg, Seth  
; TITLE OF INVENTION: Ten-35 Polypeptides and Polynucleotides and their Methods of Use  
; FILE REFERENCE: Attorney Docket No. Cura 967  
; CURRENT APPLICATION NUMBER: US/11/096,051  
; CURRENT FILING DATE: 2005-01-30  
; PRIOR APPLICATION NUMBER: 10/038,854  
; PRIOR FILING DATE: 2001-12-31  
; PRIOR APPLICATION NUMBER: 10/455,772  
; PRIOR FILING DATE: 2003-06-04  
; PRIOR APPLICATION NUMBER: 60/557,978

QY 683 RHYSOCCLIDSGDGKFYTIHTQOYDEOQRHHTSTSVDLNGRQQTIDPKVHLSMSKSD 742  
Db 1753 -VWRFRPKEQAOQK--VVFRGKLRVNGRNLLSVDFFRTKTEKTYDHRKEL-LRAYD 1818  
Qy 743 NWGQIANTHWSYCSEKTIUDVPTLTAKQLOSNNSNNYVTGKEVTTVTPSQQPIQITFD 802  
Db 1819 TSGH--PTIWL--PSSKLMAVNVTYSSIGQIAS----IORG-----TTSEK----VYD 1860  
Db 803 EAGHLQSCHTLRDG--WDRVERKETDAI-----GOCTRYOYDNTNRYVQITLD-----849  
Qy 1861 SQCRIVS--RVEADGKTSVSYTLEKSMWLLQSQYVFEYDMDRISAITMPSVARHW 1918  
Db 850 -----GTVNRKYAPFSTDILDIRUNGISLQQTDFGLSLRITQSDGGRWVATYSA 903  
Db 1919 QTIRSIGYTRNTVNPBESNASITDYNEMEGLLI-QTARLGTS-----RV-----1962  
Qy 904 GNDQCPSTVITPDCQFINTQYQPELDDAVLQVNASNETTQFSYNTPVUTQVALLKAV--AEG 960  
Db 1963 -----LFYRQRQTRLS---EILYDSTRVFTDETAG-VLKTVNLSDG 2002  
Qy 951 QSLRPIYPPSGRLKMENTNDMKNSYUMLVGLGNGYDPLTGTOKSRDTHGRVTKD 1020  
Qy 2003 FICITRYMQRQGPL-----IDQIFRSEGDGAARDED-----YVYDNSRTRVTSQG 2048  
Qy 1021 SSITKTL-----NYDDLMRHIGS-----QVTDLIAHGMLTIVTFD--GLNREGR 1064  
Db 2049 VINETPLRIDLQYFDDISGKVBGPFGKRCVYDINOISTAVWVYTKHDAHGRIKEIQ 2108  
Qy 1065 KLCDSGSGHLDIQCWSLWLTQQLANRIVLNGVILQRTBQSYDSRNRLAQYKCDGABCID 1124  
Db 2109 EFRSLMMWITIQD--NMGRVYKREIKI-GPANTTIVAY-----EVYDVG-QLQT- 2156  
Qy 1125 KYHSIVTONFTDYGHNITACHTFADGTEDHATFKFANPPTCPQLTFRVHHTPDMDPN 1184  
Db 2157 VYLNKEKMWYNYDNLGNI-----HILNPPSSARLTF-----2188  
Qy 1185 IRLKVKAGRVINID-----NHGNTENPTVYUPLGRQ-----NGQCSVYQYD 1227  
Db 2189 --IYDLDURBIRTRLGDVYRLDDEGFLRQRTGEPEYSKGLLTRVYSKGSMWIVYRD 2246  
Qy 1228 PINLVLVQHT-----DTLDCELY--RETMLVNEVRAGE 1259  
Db 2247 GLGRVRSKTSILSGLQHQLFFYADLTYPTRTHVNHSSBTSUYDYLQHLPAMEISGD 2306  
Qy 1260 MIRLRTGHTIAQORASKVLLTQDSDOSVILTSKDNQLEQASYVK--HKSTANDA 1317  
Db 2307 -----EFYIASDN-----TGT--PLAVFSSNGLMLKQTOYTAGEIYFDSNVDQ 2349  
Qy 1318 SILCYNGERADPVGSGVTHGNGYRSYDPTMLRFHTPD 1354  
Db 2350 LVIGFHGGYDPLTKLHFGE--RDYDILAGRWTDP 2384

PRIOR FILING DATE: 2004-03-30  
 NUMBER OF SEQ ID NOS: 38  
 SOFTWARE: Curaseqlist version 0.1  
 SEQ ID NO: 4  
 LENGTH: 2376  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-11-096-051-4

Query Match 3.2%; Score 283; DB 7; Length 2376;  
 Best Local Similarity 19.4%; Pred. No. 4 8e-10;  
 Matches 365; Conservative 255; Mismatches 637; Indels 626; Gaps 97;

QY 18 NEFFTOQANNFTSAVGGVGDPPRTGLYNIQITLGHVGNGNMGPTPLTLSYSPLAKTDGF 77  
 806 NQFISQPPVVSSIMG--NGRRSISCFCSCNGQADGNKULAPVA-----LAC 850

QY 78 GI-----GFEFLSVDRKNS--SLLS-----TGENVKYE 107  
 851 GIDASLYVGVDFNVYRRIFFPSGNVTSVLELRNKOFRHSSNPARYLAATPVGDLY--VS 908

QY 108 TDKTVKLQOKKDLRPERKDLCENCKYRIRHKGDIELVLTGNNNA--FD-----LKV 157  
 909 DTWTRRIVRPK--SLTGADKDTKA-----EVAGTGEQCLPDEAKCGDGKAV 956

QY 218 GYTERLEINRQNSIHNFSGLHENPLTMSSGYPIGNIGLGQWLSMTAGLGKETN 277  
 1000 -----LTSARPLCD--TSMHSISQRLEMPTD-AINPMDSIY 1035

QY 278 YSNNNQGHFPQSANLPLVPPVTLMKQPG---AQPAIAQAYSYISHNYVGGSGNIM 333  
 957 BATLMSPKOMA-----DKNGLIYFVDG--TWIRKVDONGTISTL---GSND 999

Db 1036 VLDNNVWQIQTENQVRAAERPHCOPGVPEVPGKAVOT--TLSATAAVSYGV- 1092

QY 334 NKKUDNLQGLMTEVNYGSTRESRYKDKEGHDQTVRIERTYNNHLLTSBCKON----- 387  
 1093 -----LY-ITEDEKKINRIRQVTDGEISLVA-----GIPSECDCKNDANCDC 1135

QY 388 -----GYIOTTE-----TAXAIGH-----NR-----D 406  
 1136 YQSGDGYAKDAKUSAPSILAAASPDTLYIADLGNNIRRAVSKRPLNSMNFYEVASPTD 1195

QY 407 SQPSQFQLPKTKETWRSADNSY-----RSBITETPDESQNPUTKVKD--KKTQK 456  
 1196 QELXIFDINGTHQXTVLSTGDXLYNESFSNNDIATV-DMSGNTL-RIRUPNRMVR 1253

QY 457 LISRSTH--WEVYPPAGEVDNCPPPEPYGFRFKKIIQTPYDSBFKODPEKFIQYRSLI 514  
 1254 VVSFDNQVWLTTGNGCJMSMTAQL-----EVLFTH- 1288

QY 515 GSOSHVTLKIERYISATOLINSLFLQNTDKSELGLRLKQTC-----KGENGKTY 568  
 1289 GNSSLALATKSDETGW-----TTFEDYUSE----GRJNTVTPGTGVVNLAHDMDKAIT 1337

QY 569 VWHKFTYTKQDDTLOQSHSITHDMF-TIHRQSRSRTGRLFSSTDKDIVTOMSDKL 627  
 1338 VDEBSSRREEDVSI--TSMLSIDSFSYTMVODLRSN-----QIGTGS 1380

QY 628 GRILLRTLNLNGSTPVANTLT--YDVELNNIQDNRPPFV--ITTDVNGQJNEFDGAG 682  
 1381 LRLI-----YASGLDHSYQTEPHVLAGTANPTVAKGRNMWLPGENQNT----- 1423

QY 683 RHYSOQLKSDGDOCKPFYHTQFQDEQSRHHTSTSVDYLNGSQQTDDKVLHSMKSDY 742  
 1424 --VWFRFRKEQAOQSK--VNFGRKURVNGRNLSVDFDRITKTEKYDDHFKL-LRJAYD 1479

QY 743 NWGQIANTHMSYGSEKITVDPITLTAKOLQNSNNVYQGKFTVYPSQQPQIQTIFD 802  
 1480 TSGH--PTWLL--PSKLMVNVTYSSCQIAS---TORG-----TSEK---VYD 1521

QY 803 BAGHLQSCHTLTDG--WDRVKETDAI-----GCQTIQDYNRVIQITLP----- 849  
 Do 1522 GOCRIVS--RVFDGKTMWSYTLKSMWLLHSQROVTFDWMWDLAUSLTPSPVARTM 1579

QY 850 -----GTVNRKXAPPSTDLTDIRVNGISLGQQFFDGSLRSTSQDGGRWVAYTSA 903

QY 1580 QTIRSIGHTRNLYNPPESENASTIDTNEBGLL-QTAFGLFSR-----RV----- 1623

Do 904 GNDQCPSTVITPDQFTHYQXOPPELDAVLOVASNETITQFSPSYNPVGTGALLKAV--AEG 960

QY 1624 -----LFRYRROTRLS---EYDSTRVSFRYDETAG-VLKVNLIQSDG 1663

Do 961 QSLTPPIPPSGRKMEMINDKGKSMVWTGLENQYDQGKISRDYGRVQIKD 1020

QY 1664 ECTIRTQIGPL-----LDRQIFERSBGDNMRFD-----YSYDASFRVTSMQ 1709

Do 1021 SSITKTL-----NYDLMRHIGS-----QVTDLATGMHMTTVEFD--GLNREIGR 1064

QY 1710 VINETPLPBDLYQFDDISGKVHQFGKCVVYIDQINQISTAVMTYTKHFDAHGRIKEIYQ 1769

QY 1065 KLCDSSGTTLDIQQSWMKTQOLANRVTKLNGYQRTQCYSTSRSRNRJNQYKCDGAKCPFD 1124

Do 1770 ETRSLWYMITQYD--NMGRVYKREIKI--GFPANTTLYAY-----EYDVQG-QLQT- 1817

QY 1185 IRUKYDKAGRVINID-----NHGNTENFTYDTLGLQ-----NGQGSVYGYD 1227

Do 1850 --LRYDLDRDRITRIGDQVYRQRLDEGFLRQGTEBIFEYSKGLTRVSKGSWVNTIYRD 1907

QY 1228 PLMLRVLSKT-----DTLDCBLYY--RETMVUNVERGE 1259

Do 1908 GLCRRVSKTSLSQHLOFVADLYTPTRITHVNHSSSETSLVYDQGLHFAMEISGGD 1967

QY 1260 MIRLRTQETTIAQQRASKVLLGTDSQSQVILTSQKLNLSQEASYSKG--HKSTANDA 1317

QY 1988 -----EFYIASDN-----TGT-----PLAVFSSNQMLKQIOYTAYGETYFDSDNIDFQ 2010

QY 1318 SITYNGSERADPVGVTHLGNGYRSYPPTLMFHTPD-----SISPGF----- 1360

Do 2011 LVIGFHGGLYDPIITKUHFG--RDYVILAGRWTPDEBIWRIGKDPAPPFLYMPANN 2068

QY 1361 -----AGCINPYSYCYGDPINSD--PSGH--SWOAWTGICGI 1396

Do 2069 PASKHDVRYTDVNSWLVTFGFLHNAIPGPVPKVPLTFSYELVKSQOQDDI--PPI 2127

QY 1397 AGLITATGGMIALAAGGIAIAS-----TSTTALAFGALVTS-----I 1439

Do 2128 FGYQOQVRQAKFLSLGKMARVQVSERRAGGAQSQWLFATVKSLSIGKGMLAVSQRVQ 2187

QY 1440 TSTVSGALADEASPKASSTLG-----WVNGMGAGLAESAIKGCT 1480

Do 2108 TNVNTIANEDCI-KVAUAVUNAFLYENHFTEGKOTHYFIFTTPSPSDIGLRLISGRK 2246

QY 1481 LATHLGAFAEDGENALGSTSSRSRKGWV-----SLDRBIVR 1520

Do 2247 AL-----ENGINVTVSQTSTVVNGRTRPADVEMQFGALAHVRYGMVLDDEBKAR 2296

QY 1551 NERGQVQIDHSKSGYDPMFGKOBQALVHGDXDGFPLYTHEGMKH---GK-----GPFY 1571

Do 2297 ILBOAROPALARWA-----REQBQRYDGERGARLN-WTEGKRQQLSAGKVQGYDGYV 2349

QY 1572 RHTPE--OLVYDVKDNNTVLDQ 1592

Do 2350 LSVEQYPLAD--SANNIQFRLQ 2370

; GENERAL INFORMATION:  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: MacLachlan, Timothy K  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Vernet, Corine  
; APPLICANT: Ettenberg, Seth  
; TITLE OF INVENTION: Ten-M3 Polypeptides and Polynucleotides and their Methods of Use  
; FILE REFERENCE: Attorney Docket No. Cura 967  
; CURRENT APPLICATION NUMBER: US11/096,051  
; CURRENT FILING DATE: 2005-03-30  
; PRIOR APPLICATION NUMBER: 10/018,854  
; PRIOR FILING DATE: 2001-12-31  
; PRIOR APPLICATION NUMBER: 10/455,772  
; PRIOR FILING DATE: 2003-06-04  
; PRIOR APPLICATION NUMBER: 60/557,978  
; PRIOR FILING DATE: 2004-03-30  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: CurateSeqList version 0.1  
; SEQ ID NO: 2  
; LENGTH: 2715  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-11-096-051-2

Query Match 3.2%; Score 283; DB 7; Length 2715;

Best Local Similarity 19.4%; Freq. No. 5.9e-10;

Matches 365; Conservative 255; Mismatches 637; Indels 626; Gaps 97;

OY

18 NEFFTOANNFTSAVGGVDDPRTGLYNIQITGHTIVGNGLNGLPTPLTYSPLNKTDGF 77

OY

1145 NQFISQQPPVSSIMG--NGRRRSTCPSNCQADGNKULLAPVA-----LAC 1189

OY

78 GI-----GPNFLSVPYDRN--SLISLS-----TGENKYIE 107

OY

1190 GIDGSILVGDENYVRRIFPSGVNTSBLRNKDFRHSSNPARYYLATDPVTGDLV--VS 1247

OY

108 TDKTVKLQOKKLNRFEKOKIKENYTRIHKSGDIEVLTGCFNNNA--FD-----LKV 157

OY

1248 DTNTTRIYRPK--SLTGAKDTKA-----EVVAGTECOPFLDEARCSDGGKAV 1295

OY

158 PKKLINPAGHAIYIDWNFEATQPLRRIYDLDGHIDPLNLQEVQLKITLTLPQKE 217

OY

1296 EATLNSPKGMAV-----DNLGLYFVDG--TMIRKDONGIISP-----GSND 1338

OY

218 CYTERLFLNRLNOLNSTHNFSIGNENPLTWGSGTYPIGKONGILGOMITSWMTAPGGKETVN 277

OY

1339 -----LTSARPLTC-----TSHHSQVRLEPTDL-AINPMNDSY 1374

OY

278 YSNNNQGHHRPFPQSANLPUVPLPVTLMKQPG---AGOPAQAENSYTSRHYVGCGSGNLTW 333

OY

1375 VLDDNNVTLQTENRQVIAMGRPMHCQVPGVEPVPGKHAVOT--TLESATAIAVSYGV- 1431

OY

334 NINKLNDNLYGLMTEYNGESTERRYKODKEHQDIOVIRERTINYHILTSEQQON----- 387

OY

1432 -----LY--TETDEKKKINRQVTDGEISLA-----GIPSECDCKDANCDC 1474

OY

388 -----GYIQTRE-----TAYVATIGH-----NF-----D 406

OY

1475 YQSGDGQYAKDNLKSAFSSLAASPDGTLYIADLGNIKIRAVSKNKLPLNSKMFYEVASPTD 1534

OY

407 SOPSOPOLPKXKTETRSADSY-----RSELTETTDEDESGLPLTKVLD-KKTOK 456

OY

1535 QBLYIJDINGTHQYTSLVTDYLYNPSYSNDTIAVT-DSNGNLT-RIRRDPNMPV 1592

OY

457 LISPSIH--WYYPAGEVDNDCCPPEYFTRVKLIQTYDSEKFDPKFIQYRSLI 514

OY

1593 TVSPDNQVWILITIGTNGCLKNTAOGI-----ELVLFYH-- 1627

OY

515 OSOSHWTLKIBERHYATQOLNASTLROQNTKSELGRLLKOTE-----KGENGKTY 568

OY

1628 GNSGLLATKSDEGW-----TTFRDYDSR-----GRLTNTWPQGVVNLHGMDKATT 1676

OY

569 VVHKFTVTKQDITLQQSHSITHDNP-TIHSQVRSRYTGRFLFSQDUDTQVTONSYDKU 627

OY

US-11-096-051-2

Db 1677 VDISSSRBVEDSI--TNSLSSIDSFTVMQDOLRNSY-----QIGYDGS 1719

OY

628 GRILLTRIANSGTYANTIE--YDEVELNLNQDDNRPPFV--ITTDVNGNQLNEFDGAG 682

Db

1720 LRIT-----YASGLDSHOTEPPLVLAGTANPTVAKRNMLPGENGQN----- 1762

OY

683 RHYSOCLADSQDGCKFTHTQOYDECGRHHTSTSVDLTSNQROTDQDDKVLHSMSSYD 742

Db

1763 --VWRFRKEQAGK--VNVFGRKLRVNGRNLLSVDFTRKTKEVYDHRKEL-LRIVD 1818

OY

743 NWGQIANTHWSYGSEKTIUDPITLTAKOLOSNSNNNTGKRVTTYPSQOPIOTIFD 802

Db

1819 TSGH--PTWL--PSSKUAMAVNTYSSITQGIAS-----TORG-----TSEK-----VIDD 1860

OY

803 EAGLHQSCHTLDRG--WDRVRKETDAI-----GQCTIYQDYNVRVQITLD----- 849

Db

1861 GOERIVS--RVEADGKTVSYTYLKRKSMVLLHSORQYFEYDMDRUSAITMSVARHTM 1918

OY

850 -----GTVUNRKYAPFSTDLTDIRUNGISIGQQTIDGLSRLTQSODGGRWAWTSA 903

Db

1919 QTIRSIGYRNTNPESMASITDYNREGULL--QTAALGTS-----RV----- 1962

OY

904 GNDOPCPSTWITPDCOFITQYQFBLDDAVLOQASNSETIQFSNPVPTVALLKAV--AEG 960

Db

1963 -----LFIYRQRQLIS---EILYDSTRVSFTIDETAG--VLKTVNLSQDG 2002

OY

961 QSLTPPIVPSGRKXMENTNDKMSVYLMRLGYLENGVDTGTLGIQKSRDTHGRVTOQKD 1020

Db

2003 FICTIYRIGPL-----IDQIFRSEGDVMNARDED-----SYDNDSPRVTISKQ 2048

OY

1021 SSIKTL-----NYDDLNKHGS-----QVTDLAIGHMLTTVFD--GLAREGR 1064

Db

2049 VINSTPLPILDYQPDDGSKVQFGKFRVIVYDINOISTAVMVYTKHDFASHKRIKE-TQY 2108

OY

1065 KLCDSGSHFDLIDOCQSWLQYQQLANRIVKLNGVQRTQSYDSRNRLNQYKCDGAECPTD 1124

Db

2109 EIFRSLMMWITQD--NMGRVTKREKI--GPFANTTIVAY-----BYDVDC-QLQT- 2156

OY

1125 KYGHSIVTONFTDYGINTACTHTFAAGTEDIATPKFANPTPDCQLTQEVHHTPDMDDN 1184

Db

2157 VYLMKIMWRYNVDNLN-----HILNPSNSARLT----- 2188

OY

1185 IRJKVDAGRVINID-----NGANTENPTVYTLGRLO-----NGQCSVYQD 1227

Db

2189 --LRYDRLURTRITRGLDQYRDLDEGFLFQRTFLEFEYSSKGULTRVSKSGMWIVYRYD 2246

OY

1228 PLNRLVSKT-----DTDCCELY--RETMLVNEVRNC 1259

Db

2247 GLGRRVSKTSLSGHLQFYADLTYPTIRTHVNHSSBRITSLYDQHLFAMEISGD 2306

OY

1260 MIRJRTGETTIAQRAKSVLQLTGSQDOSVILSKDQNLSBQAYSAVCK--HKSTANDA 1317

Db

2307 -----EYIADN-----TGT-----PLAVISSNGMLQYQTAEGIYFDNSIDQ 2349

OY

1318 SILGYNGERADPYSGVTHGNGYTSYDPLMRFPTD-----SLSPPFG----- 1360

Db

2350 LVIIGFPHGSLYDPLTKLHFRGE--RDUYDILAGRWTPTDIBWKRIGKDAPPFLYMRNN 2407

OY

1361 -----AGJNPYSCVLGDDPNRSD--PSGHL--SWQAWTIGIGMI 1396

Db

2408 PASKIHDVKDQYITDVSMLVTFGLHNAIPGFVPKDLTESYLVQSQWODI--PPI 2466

OY

1397 AGLILITATGGMMAAAAGIAATAS-----TSPALRAGLSVMSD--I 1439

Db

2467 FGVOQOVARAKATLSGKMAEVOSRRAGGAOSWLNPATVUSLIGKGMALVSQGRVQ 2526

OY

1440 TSIVGALADASPQASSLG-----WISMGMGAGLAESAIKGTR 1480

Db

2527 TNVNTIANEDCI--KVAANVAFYLENLIFTIEDKTHYFIKTTPESDQGLTRLUTSGR 2585

OY

1481 LATHGAFAEDEGENALKSTSESSRIKGVTR-----SLDREIV 1520

Db 2586 AL-----ENGININTVSGSTTWNGRTERFADYEMQFGALAHJVRVGMTLDEBEK  
Qy 1521 NEBGQVIRKDHSHGTYTDNFNGKGCBQALLVHGDKGFLYHTEGNHN-----GK-----GPYT  
Db 2636 ILBQRQRALARAWA-----REQQVRDGEGERCARLN-TEGEBRQLJSAGKVQGYDGYV  
Qy 1572 RHMPE--OLVDYKDDNNIVDLTQ 1592  
Db 2689 LSVEQQYPLAD--SANNIQFLRQ 2709

RESULT 6  
US-10-821-234-1097  
Sequence 1097 Application US/10821234  
Publication No. US200505511A1  
GENERAL INFORMATION:  
APPLICANT: Labat, Ivan  
APPLICANT: Stacie-Craig, Birgit  
APPLICANT: Andarmani, Susan  
APPLICANT: Tang, Y. Tom  
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
FILE REFERENCE: 821A  
CURRENT APPLICATION NUMBER: US/10/821, 234  
CURRENT FILING DATE: 2004-04-07  
PRIOR APPLICATION NUMBER: US 60/462, 047  
PRIOR FILING DATE: 2003-04-07  
NUMBER OF SEQ ID NOS: 1704  
SOFTWARE: pt\_SEQ\_Genes Version 1.0  
SEQ ID NO 1097  
LENGTH: 1094  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-821-234-1097

; PRIOR FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: 60/208263  
 ; PRIOR FILING DATE: 2000-05-31  
 ; PRIOR APPLICATION NUMBER: 09/939398  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: 60/227800  
 ; PRIOR FILING DATE: 2000-08-25  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1609  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 170  
 ; LENGTH: 2333  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-453372-170

Query Match 3.1%; Score 272.5%; DB 6; Length 2333;  
 Best Local Similarity 19.5%; Pred. No. 2.3e-09; Mismatches 652; Indels 603; Gaps 99;  
 Matches 365; Conservative 251; Mismatches 652; Indels 603; Gaps 99;

Qy 49 GHIVGNG-NLGPILPLTISYSPLKTDIGFGIGENFGISVYDRENLSLSTIGENYK--- 104  
 Db 666 GHFLQKSFOSRPNALASTPWI--DKTDA-----YGORVYGLSDAVVSV-GREYETCP 713

Qy 105 -VIETDKVVKLQQKKLD-NL--RFEK--DLKENCYRIRHK-SGDIEVLTG--- 147  
 Db 714 SLIIVKRPVALLQDOPFEDPSNLGGWSLDKRHILNKSG---IHKGTGBNQFLNQPAII 770

Qy 148 ---FNNNAFDLKVP----KLLNIPAGHAYIDMNFEATOPRLNRYDDLDGHIP 195  
 Db 771 TSIMONGRRRSISCPSCNGLAEGNKGLLAVLGDSL----- 810

Qy 196 LNLNBYYQGLIKTIITLFPGOKEGYRTELFLNROL---NSTIRNFSLGNEPLTWSEGYT 251  
 Db 811 ----YVGDFNMRIFRIPSKRNVTITLER--NKEFKHSNNPAHKCYLA-VDPVSGSLVVS 862

Qy 252 PIGRNLGILGOWITSMTAPGGLKETTNYSNNNQHHPQSANLPVLPVYLKMQVPGAQOP 311  
 Db 863 DTNSRRIY---RVKSLSGTLAGNLSNEVWAG-----TGEQCLPP--DEARCGDCK 908

Qy 312 AIQAFYSYVSHNYVGGGSNIWN-----NKLNDLYGLMTEYNGSTSERRYK---DK 360  
 Db 909 AIDA-TLMSRGIAVDKQNLIMYFDATMIRKVHQ-NGLISL-LGSNDLTAVRPLSCD 964

Qy 361 EGHDOIVRTE-----RTVNNYHILTSBCKQONGYIQTETAYAIGHNFDSQPSOP 412  
 Db 965 SMDVAQVRLWEPTDLAVNPMDNSLXLV----ENVILRITENHOVSLA---GRPMIC 1015

Qy 413 QLPKKTTETRSADNSYRSITETTDESIGNPLTKVOKKTQKISOSTHWEXYPAGE 472  
 Db 1016 OUPGIDYSLSKLAHSALESASATSAISHTGVLYTETDECKINLRQVTINGECLLAGA 1075

Qy 473 VDNCEPE-----PYGFRFKKLIQTPYDSEBFKDP 503  
 Db 1076 ASDCOCKNDYNCYCSDGDAYATDAILNPSPLAVAPDG-TIYADLGMRTRAVSKNP 1134

Qy 504 -EKPTOYRVSLLIGSOSHVTKIEERHYSATOLLNSTL--FQNTNDSKEBLGRJLKQTEC 558  
 Db 1135 VLNAPNQYEASPGEOBLVYFNADGIQHQTSLVTCBLYNFTSTD-NDVTEL---- 1188

Qy 559 TKGENGKTYSVHFKFTYKDDTQLOSH-----SITTHNFTHRSO----- 600  
 Db 1189 -DNNENSLKJ-----RRSSGMPRHLAMPDNQNTITLTWTGNGLKVYSTONLEGLM 1239

Qy 601 VSRVYTGRLFESTDTKDIVTOMSYDKUGRLJRTNLTG-----PIANTLTVOEMLN 654  
 Db 1240 TYDGNVGLLAKTSDETGWTTHYDHEGRUINVRPTGVUTSLRHEMEKSITIDIENS 1239

Qy 655 ODDNRPPFVITTDVNG-----NQLRNEFD-----GAGRH----- 684  
 Db 1300 DDD-----VTVTINLSSVEASYTUVQDVQRANSYQOLNNGTIRVMYANGKGSIFSEPHVL 1354

Qy 685 -----VSOC-----LKQSDGDGKPYTHTQOYDEQSRHHTSTSYDVL 721

Qy 1355 AGTTPPTIGRCNISLPMENGINSIERLKEOTKGK-VTFCRKLVRHGRNLST---DYD 1411  
 Qy 722 TNGR-QDTPDKVHLMSKSKYDNGQ-----IANTHNSY-----GUSBK 759  
 Db 1412 RNRTEKYDDERFKFTARLIVYQGRFLWPLSSGLAVUNSYFFNGRLASQHORGAMSER 1471  
 Qy 760 ITVDPITLTATKOLQSNNSNNVQGKEVITYTSPSQQTQITLFDEAGHLSQCHTLRDGW 819  
 Db 1472 TDID-----KOGRTISRMFADGK-VWSYSLDKSMWLLQSQQRQY----- 1510

Qy 820 RYRKETDAIGOCHIYQVNNYNNVIOITLPDGITVNRCVAPFSTDTLTDIRVNGISHGQQ 879  
 Db 1511 -----IEFYINSSDRLLAVTIPS---VARH--SMSTHTSIGYR---NINP 1548

Qy 880 TFDCLSLRTQSODGGRWVAYTISAGNQCPSTVTPQGQFINTQY--QPELDAAVQWS 937  
 Db 1511 PESNASVIFPDYSDPGRIN-----KTSPLSGRQVQYKYGKSKUSBIVN---S 1594

Qy 938 NEITQQFSYNPVIGAL--LKAVAEQGSIIMPVPSGRKMMENINDMKKMSYLTLRGEN 995  
 Db 1515 TAVT--FOYDTEGVLKVNVLQSGGFCTIRTRKGP-----VUKQIVPFBEGMV 1645

Qy 936 GYDLTGTTQIKTSRDTKRGVTOIKDSSKTTL-----NYDDN--RHIGS-----QVT 1041  
 Db 1646 ARFDYT-----YHDNSPRIASIKPVISSTPLPVLDYRIDEISGKVERFGKCVIVYDIN 1699

Qy 1042 DLATGHMHTTVERD--GLENREGRKLGUDSSGHTLDQOSWLTQDOLANRIVYLN--GV 1096  
 Db 1700 QITTAWMILSKHDFTHRIKEYQEMRSLSMWMWYQYDSM-----GRVTRKELKGP 1753

Qy 1097 LORTBQYSYDSRNLNQXKCDGA-----ECPTDKYQHGSIVTQNFTYDIXGNITACHT 1149  
 Db 1754 YANITKTYD-----YDGSGOLOSVANDRPTWY-----SYDLNGNL----- 1791

Qy 1150 FADOTDEDHATKCFANPTDCQCLETTEVHHHHDPMDDNIRKDYKAGRVINI-----TDHNG 1203  
 Db 1792 -----HLLNPGNSVRL-----MP-----LARYDRDRTLGQVYKIDDG 1827

Qy 1204 -----NTNRFPTYDGLR-----ONGSGSVYGDPLRVLVSOKTDIDCELYRETMLV 1252  
 Db 1828 YLCORGSI-----FEYNSKGILTAYNKAWSGSVQXQYDGVERRASXTN-LGHHLQFYSLH 1886

Qy 1253 NEVER-----NGEMILU--RIGETIAQQRASKVULGTDQSOS--VILTSKDQULS 1300  
 Db 1887 NPTRITHVNHSEITSILYYDLOGHLHAWESSGGERYVASDNTGTPLAVFINGLMK 1946

Qy 1301 QEAVSAY3K--HKSTANDASILGNGERADPVSGVTHLNGRISYDPLMLRFITPD--- 1354  
 Db 1947 OLQYTAYSGIYVDSNPDMQWIGFHGJYDPLTKLVTQ--RDYDVLAGRWPSDVTW 2004

Qy 1355 -----SLSPFGAGGINPNSYCYCLGDPINNS-DPSGHL---SWQAWTG----- 1392  
 Db 2005 KNVGEKAP-----NLMFKNSNPLSSLDKKNVYDVKWSLUMFGQLSNIIPGPFR 2059

Qy 1393 -----GMGTAGLILITIATGMAIAAGG-----IAAIA----- 1421  
 Db 2060 KMYFVPPPYELSESQAENGQLTGVOQTERHNOAQMALLEGOVITKOKLHASIREKAW 2119

Qy 1422 -SITSTALAFGL-----SVTSDTISIVSAGLEDASPKASSIG----- 1459  
 Db 2120 FATTPILGKIMPAIKEGRVITCVSSA-----ED-SRKVAVANNVYLDKMHYSISBK 2175

Qy 1460 ---WVSMGNGAAGIA-----USA-----IKGGTKLATHIGAFAEDGE 1493  
 Db 2176 DTHVFKVIGSADGDLVLTGTTGRKVLSEGVNVTVSQPPLVINGRTRRTNI---EFQY 2231

Qy 1494 NALKSTSESSRIKMGVNT-RSLDREIVNNEEGQVQKDISRGYTDNFMGKGBQAITVHGDK 1552  
 Db 2232 STLLIS-----IRVGLPTDDEKAR-----VLDQARQRALGATAWAKEQOKARD-G-R 2278

Qy 1553 DGFLYHTEGNK 1563

Db 2279 EGSRILWTEGEK 2289

RESULT 8  
US-10-453-372-114  
Publication No. US20060003323A1  
GENERAL INFORMATION:  
APPLICANT: Alsobrook, et al.  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEAR ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-589 A  
CURRENT APPLICATION NUMBER: US10/453,372  
CURRENT FILING DATE: 2003-06-03  
PRIOR APPLICATION NUMBER: 09/789390  
PRIOR FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: 60/185967  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 09/23187  
PRIOR FILING DATE: 2001-03-29  
PRIOR APPLICATION NUMBER: 09/195792  
PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: 09/839446  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 60/199476  
PRIOR FILING DATE: 2000-03-25  
PRIOR APPLICATION NUMBER: 09/863776  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: 60/208263  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 09/939398  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 60/227800  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1609  
SOFTWARE: Cursequist version 0.1  
SEQ ID NO 114  
LENGTH: 2662  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-453-372-114

Query Match 3.1%; Score 272.5; DB 6; Length 2662;  
Best Local Similarity 19.5%; Pred. 2.8e-09; Mismatches 365; Conservative 251; Mismatches 652; Indels 603; Gaps 99;

Qy 49 GHIVGNG-NLGPPTPLTYSPLNKTDI[GFGIGFNGF]SVYDRKNSLISLTG[ENY]-- 104  
Db 995 GHIFQKSFOASPNLASTFW---DKTDA-----YGQRVYGLSDAVSV--GHEYCPC 1042

Qy 105 --VIETKTVKLOOKKLD--NL--RPEK--DLKENCYTRIHK-SEDIEVNG---- 147  
Db 1043 SLIWKEKRTALLOGLFELDPSNLSGWSLDKHKHILWKG---IHKGTGENQPLTQPAIT 1099

Qy 148 --FNNAPDLPKP---KKLN[PAGHAYIDMANFEATOPRLRIYDDGDHP 195  
Db 1100 TSIMNGNRHSRISCPSCNGLAEKGKLAVVALVGIDESL----- 1139

Qy 196 LNLYEQGLIKTIITLFLPGQKEGYTERELPFLNROL---NSIRNFSLGNENPILWFCYT 251  
Db 1140 ---YVGDFPNVIRIFPRBNNTSILER--NKEPKHNSNPANKYLA-VDPVSGSLXVS 1191

Qy 252 PIGKNGILGQWMTSMTAPGGLKETWNYSNNQGHFPOSANLPTLPVYTLMKVPGACP 311  
Db 1192 DTMNSRRYI---RVKSLSGTKOLAGNSEVAG-----TGBOCCLPP---DEARCGLDGK 1237

Qy 312 A[Q]EYEVTSVSHNYVGGSGUINN-----NKDNLYGMLTMEYMGSTESRRIK---DK 360  
Db 1238 AIDA--TMSMRGIAVDKNDGMVYFDATMIRKQDQ-NGISTL-LGSNDLTAVRPLSCDS 1293

Qy 361 EGHQIVRIB-----RTNNYHILLSECKQKQYOTTEAYAYAIGHNFDQSQF 412  
Db 1294 SMDVAQVRLWEPPTLAVNPMNDNSLYL-----ENVILRITENHQVSTIA---GRPHIC 1344

Qy 413 QLPKPKETETRSADNSVRESEITSTPDESIGNPLTKVVKDKKQK1ISPSTHWYYPAGE 472  
Db 1345 OVPGDYDISLSKLAHSASALSAATISHINGVLYTETDBKCKNRLROVTINGETCLLAGA 1404

Qy 473 VDNKPPE----- 473 VDNKPPE-----  
Db 1405 ASDECDKODVNCCNCYSGDDAYATDAILNSPSSLAVAPG-TIVIADGCRITAVSKNP 1463

Qy 504 --EKPIQKRYSLTGQSQSHNTKLEERHYSATQNLNSTL--FOYNTDSELGRLKQTC 558  
Db 1464 VLNHNQYEAASPOEQLKVFENADGIKTYVSLVGETLKNFTSTD-ADVTEL---- 1517

Qy 559 TKGNGKTVSWKHPFTVKQDPTQOSH-----SITHDNFTTHRSD----- 600  
Db 1518 --DNNGNSLKI-----RDSSGMPRLHMPDNOITLTGNGGLKVUSTOMELGJM 1568

Qy 601 VRSKTYGKRSFSDTDKOTVOMSYDKAHLUTLNST-----PIANTLTQHML 654  
Db 655 QDDNRPPFVTTDONG-----NOLNEED-----GAGRH----- 684

Db 1629 DDD-----VTWITLSSVEASYTVQDQVANSQLCNGTIRMYANGMISHTSEPVIL 1683

Qy 685 -----VSOC----- 685 -----VSOC-----  
Db 1684 AGTITPTIGRCNTSLPMENGNSIEWRLKQEIKGK-WTFGRKLVRGGRNLSTI-DYD 1740

Qy 722 TNGR-QQDPDKHLSMSKSYDWMQ-----LANTWMS----- 759  
Db 1741 RNRTEKTYDHDHFKPTLHITYDQYGRPTFLPESGLAAVNVSYDNGLAGLORGAMSER 1800

Qy 760 ITDPIPITATKOLQSNNSNVQTKCEVTTYTPSQQPIOITLFDABGHIQSCHTLTRDGWD 819

Qy 772 TDD-----KOCRIVSMFAOK-TVSYSVLSKMSWMLQDQY----- 1839  
Db 1801 TDID----- 1801 TDID-----  
Qy 820 RVRKETDAIGQCTIYQVQNYNRYVIOITPDGTUNRKAPFSTDLTDIRVNGISLGQ 879  
Db 1840 -----I[F]BYBSSDRLLAVTIPS--VARH--SMSTATISGYR---NIVNP 1877

Qy 880 TFOGLSRJQSQDGGVWAVYTSAGNDQCPSTVTPDQSPRIQY- QPBDLAVIYAS 937  
Db 1878 PESNASVIFTDYSDDGRIL-----KYSFLGTRQVFKYKSKLSLVYD--S 1923

Qy 938 NEITQOFSKYNPVTGAL--LKAVARGOSTPPIYPSGRKJHENINDMKCNSYMLTGLJEN 995  
Db 996 GYDULGTGIOKTSRDTHRVTOKDSSKTTL-----NYDDLN---RHGS-----QVT 1041  
Qy 1975 ARFDY-----YHDNSPRIASTKPVISBTPLVUDLYRDEISGVFVHFGKFGVYIDIN 2028

Qy 1924 TAVT--FQYDTEGVLKVNLLQGGSFSTIRYKGP-----VHQIYPRSEGMN 1974  
Db 1996 GYDULGTGIOKTSRDTHRVTOKDSSKTTL-----NYDDLN---RHGS-----QVT 1041  
Qy 2029 QITTAUTMILSKHDFTHRKEVYEMFRSLWMTWQYDSM-----GRVIREKLKGP 2082

Qy 1042 DLATGMMTTTVERD--GUNRGRKLGDDSGHTLDQOSWIKTQLANRIVKL--GV 1096  
Db 1097 LORTEQKSYDSRNRNLQKCDGA-----ECPTDKGKHSIVQNFTDIYGNITACTTT 1149

Qy 2033 YANTKTYD-----YDGDGQLOSLVWDRFTWY-----SDINGNL----- 2120

Qy 1150 FADQTEDHATFKFANPPTCPOLTEVHHTPDMDNTRIKYKAGRVINI-----TDNG 1203  
Db 2121 -----HLLNPGNSVRL-----MP-----LRYDLDRITRGDWQYKIDDG 2156

Qy 1204 -----NTNFPTVTLGR-----ONGEGSGVYQDPLRVLVQSKTDTDCELYRETMLV 1252  
Db 2157 YLCQRGSDIPEYENSKGLTTRAVNKAQSGVQVRYDGVGRASRQTN-LGHHLQFYDSLH 2215

Qy 1253 NEVR-----NGEMIRL--RTGETIAQQRASKVLTGDSQS--VILTSKQML 1300  
Db 2216 NPTRITHYHNSSEITSILYDQHFAWESSGEEYYVADNTGIPPLAVISINGLMIK 2275

QY 1301 QEAYSYGK--HKGSTANDASILGYNGGERADPVSGVTHINGRYSVDPYLMLRHPTD---- 1354  
Db 2276 QLQYTYGBIYDSDNPDFOMVTPGPHGLYDPLTKLVHTQ---RDYDVLAGRTWSPDYTMW 2333  
Qy 1355 ----SLPPFGAGGINPNSYCLGDPNTS-DPSGHL---SMOWATGI----- 1392  
Db 2334 KNYKEKAPP----NLYMFKSNPNPLSSEBLDNKVNIVTUVKSMWVMPGQLSNVPGFPR 2388  
Qy 1393 -----GMGTAGLLTTATGCMATAAGG----IAAIA----- 1421  
Db 2389 KMYFVPPPYELSESQASENGOLITGVQQTTERHQAFMALEGQVITKKLHASIREKAGHW 2448  
Qy 1422 -SISTALAFGL----SVTSDTISVSGALEDAKASSILG----- 1459  
Db 2449 FATTTPICKGIMPAIKSGRVTGVSAS--ED-SIRKAVTANAYLDKMHKEYSTEGK 2504  
Qy 1460 ---WVSMENGAAAGA----ESA-----TKGGTKULATHIGAFABEDGE 1493  
Db 2505 DTHYFVKIGSADGLVLTGTTIGRKVLESGVNNTVSQTLIVINGRTRRFINI--EFQY 2560  
Qy 1494 NALIKSTSESSRIKGVN-RSLPREIVNEEGVQVTDHSRGYTNFMCKGEPAAILVHGDK 1552  
Db 2561 STULLS-----TRYGLPTDLDDEKAR-----VLDQRQRALGTAWAKEQOKARDG-R 2607  
Qy 1553 DGFLYKHTEGNK 1563  
Db 2608 EGSRLIWTEGK 2618

RESULT 9  
US-10-453-372-148  
Sequence 148, Application US/0453372  
GENERAL INFORMATION:  
Applicant: Alsbrook, et al.  
Title of Invention: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
File Reference: 21402-589\_A  
Current Application Number: US/10/453,372  
Prior Filing Date: 2003-06-03  
Prior Application Number: 09/789390  
Prior Filing Date: 2001-02-23  
Prior Application Number: 60/185967  
Prior Filing Date: 2000-03-01  
Prior Application Number: 09/823187  
Prior Filing Date: 2001-03-29  
Prior Application Number: 60/195792  
Prior Filing Date: 2000-03-10  
Prior Application Number: 09/839446  
Prior Filing Date: 2001-03-19  
Prior Application Number: 60/199476  
Prior Filing Date: 2000-03-25  
Prior Application Number: 09/863776  
Prior Filing Date: 2001-05-23  
Prior Application Number: 60/208263  
Prior Filing Date: 2000-05-31  
Prior Application Number: 09/939398  
Prior Filing Date: 2001-08-24  
Prior Application Number: 60/227800  
Prior Filing Date: 2000-08-25  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1609  
SOFTWARE: CuraseqList version 0.1  
SEQ ID NO: 148  
LENGTH: 2724

Qy 1057 GILFOXFSQASPLNASTPIW---DKIDA-----YQORTVIGLSDAVSV-GFEETCP 1104  
Db 105 --VIETDKTVKUQOKUD--NU--RFEK---DLKENCYTRIHK-SGDIVTLG---- 147  
Qy 1105 SHLWKRKTALLQFELDPSNNGWSLDKHLNLNVKG---ILHKGNGENQPLTOOPAI 1161  
Db 1162 TSIMNGRRRSSTCPSGNGLABGNKLAQVALAVGDGSL----- 1201  
Qy 1196 LMLBEYQGLIKITYLTLPQKREGYTRBLFLRFLRQL---NSIHNSIGNENLTWSRGYT 251  
Db 1202 ---YWGDFNVIIRRIFPSRNVTSLR--NKFPSNNPAPKYL-A-VDVGSLNVS 1253  
Qy 1252 PCKGNGLQGQMTSMAPGQKJGETVWVSNNGQHHFQOSANUPVLYVTLKQVPGAGQ 311  
Db 1254 D'NNSRRIY----RVKSLSGTRKDLAGNSEWVAG----TCBQCLLR-----DEARCSDGK 1299  
Qy 1312 AJOAEVYVTSHVNVGGNSGMIN-----NLKDLNGLMTEFYNSTESRKY---DK 360  
Db 1300 AICDA-TLMSPRGIAVDKNGLMLMFVDMATMKUDQ-NGISTL-LGSNDLTAVRPLSCDS 1355  
Qy 1361 BHDQIVRIE----RTVNYHILITSECKHQGNYIQTETAYAIIIGHNFDSPSQF 412  
Db 1356 SMDVAQVRLEWPTDLANPMONSLYVL---ENNVLIRTEHNHOVIA---GRPHC 1406  
Qy 413 QPKTKTBWRSADNSRSETETTDESGLPLKVKDQKQKLISPSTHVEYYPAGE 472  
Db 1407 QPGIDYSLSKLAHSALSAASAIASHTGVWYIITERDEBKKINRLRQVTTBCILLGA 1466  
Qy 473 VDNCPPE-----PYGFPRVKKLQTPDSEFKDP 503  
Db 1467 ASDCDCRNDVNNGNCYSSDDAVYDATLINSPPSLAVADG-TIYADGNIRRAVSGNKP 1525  
Qy 504 --EKFIQYRYSLIGSOSHUTLKIEERHYSATOLLNSTL---FOYNTDKSBLGRLLKOTEC 558  
Db 1526 VIANAFNOYEASAPGEQELYVERADGHQYTVSLVTFBLYNFTYSD-NDVIELI----- 1579  
Qy 559 TKGENGKTYVSHKFTVTKQDDTLQOSH-----SITHDWTIHSQ----- 600  
Db 1580 --DNNGNHSKI-----RROSMPMRHLMPMDNOIILTWTGNGLKVUSTONLBGLM 1630  
Qy 601 VRSRYTGRFLSFSDTKDQVTONSYDKGQLLRTLNGT-----PYANTLTVDJNL 654  
Db 1631 TTYGNTGILATSDDETGWTFPDYHGRGLTNTVTRPGVUTSLREMEKSITDIENS 1690  
Qy 655 QDNRPPPEVITTDUNG-----NOLRNFP-----GACRH----- 684  
Db 1691 DDD-----VTVINLSSVEASYTVVQDQVRSYOLCNGTLRVMYANGMGTSPHSBPHV 1745  
Qy 685 -----VSQC-----LKDSDGDKGYTIHQQYDPROGRHITSTSYDL 721  
Db 1746 AGSIITPTIGRCNISLPMENGLNSIEMRKEQIKGK-VTICKRLVHGRLNLSI-DYD 1802  
Qy 722 TNGR-QQFDPKHLSMSKSYNMGQ-----TANTHRY-----GSEK 759  
Db 1803 RNIRTEKYYDDRKFKTRIYIYQVGRPLWPLSSGLAVNVSYFFNGLRAGIORGAMSER 1862  
Qy 760 ITWDPIITATKOLQSQUNNVNGKETTYTTSQSQPIQITLDEAGLQLQSCITLTDQGD 819  
Db 1863 TDID-----KQGRIVSRMFDGK-WMSYSVLDKSMVLLQSOQY----- 1901  
Qy 820 RVRKETDAIGQCTIYQDNYNRYVIQITLPDGTIVNRKVPFSTDLTDIRNGISLCQ 879  
Db 1902 -----IPEVDSSSLALVTPS---VARH--SMSHTSIVR---NTYNP 1939  
US-10-453-372-148

1986 TAVT-~~RAYDTEGIVKVNKNLQGGGFCSTIRYKGPL~~-----VKQIYRFSEEGMVN 2036  
 PROR APPLICATION NUMBER: 09/839446  
 PROR FILING DATE: 2001-03-19  
 PROR APPLICATION NUMBER: 60/199476  
 PROR FILING DATE: 2000-03-25  
 PROR APPLICATION NUMBER: 09/863776  
 PROR FILING DATE: 2001-05-23  
 PROR APPLICATION NUMBER: 60/208263  
 PROR FILING DATE: 2000-05-31  
 PROR APPLICATION NUMBER: 09/939398  
 PROR FILING DATE: 2001-08-24  
 PROR APPLICATION NUMBER: 60/227800  
 PROR FILING DATE: 2000-08-25  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 1609  
 SOFTWARE: Curaseqist version 0.1  
 SEQ ID NO 136  
 LENGTH: 2733  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-453-372-136  
 Query Match 3.1%; Score 272.5; DB 6; Length 2733;  
 Best Local Similarity 19.5%; Pred. No. 3e-09; Mismatches 652; Indels 603; Gaps 99;  
 Matches 365; Conservative 251; Mismatches 652; Indels 603; Gaps 99;  
 Qy 49 GRHVNG-NLGPTPLTYSPLANTDGFGLFNFGLSVWYRKNSLSLSCENYK-- 104  
 Db 1066 GHFOKSFOASPNIASTFW--DKTDA----YGORVYGLSDAVVSV-GPEYETCP 1113  
 Qy 105 -VIETDKTVKLOOKKD--NL--REFK--DLKENCYRILHK-SGDIEVITG--- 147  
 Db 1114 SLTLMWERKTALLQFELDPSNLGGWSLXKHHLINVKSG--1LHKGTGENQFLTQQPAII 1170  
 Qy 148 --FNNNAFDLKVP----KKLNLPAGHAYIDVNEATOPRLNRIVDDLGHDIP 195  
 Db 1171 TSIMGNGRRSISCPSCNGLAEGNKLLAPVALAVGDSL-----1210  
 Qy 196 LAMURYQSLIKITLTPLPGQKCYRTELFLRQL---NSTHNSLGENENPJTWSGFT 251  
 Db 1211 --YVGDPNVYTRIFPSRNNTSILER--NKGKFHSNNPAHYKYL-VDPVGSLYS 1262  
 Qy 252 PIGKNGLQWMTSMTAGLKLQVTSNNNOGHHPFOSANLPVLPYTLMKQVPGAGOP 311  
 Db 1263 DTNSRRIV---RVKSUGTKDLAGNSEWAG----TGSCCLP---DEARCQDGK 3308  
 Qy 312 AICQBYSYTSHNTYGGSSGMINN-----NUKUDNLGYLMTEVNGYSTESRVRK---DK 360  
 Db 1309 AIDR-TLMSPRGIAVDKNGLMMFVDAWMIRKVDQ-NGLISTL-LGSNDLTAVRPLSCDS 1364  
 Qy 361 EGDHQIVRE-----RTYNNPHLSECKQONGYQTTTAYAIGHNHDQSQF 412  
 Db 1365 SMVDAQVRLBWPFLAVNPMDMSLYVL---BNVNLIRTHENQVSA---GRPMHC 1415  
 Qy 413 QLKTKYIETWRSADNSTSEIBETTFDESGNPILKVIKKOIKISPSTHWYYPAGE 472  
 Db 1416 QVPGIDYLSLKAHLAISEASAIALAISHTGVUYITRDEKKINRLQVTTNBICLLAGA 1475  
 Qy 473 VDNCPPE-----PYGFTFRVKKLTQTPDSEFKODP 503  
 Db 1476 ASPECCKNDVNKNCGYSGDDAYATDAILNPSPLAVAPDG-TIYADLGNIRKRAVSKNP 1534  
 Qy 504 -EKFIOYRSLIGSQSHVTLKIEERHYSATQDINSLU--POYNTOKSELARLIKQTEC 558  
 FILE REFERENCE: 2140-589\_A  
 CURRENT APPLICATION NUMBER: US/10/453,372  
 CURRENT FILING DATE: 2003-06-03  
 PRIOR APPLICATION NUMBER: 09/789390  
 PRIOR FILING DATE: 2001-02-23  
 PRIOR APPLICATION NUMBER: 60/185967  
 PRIOR FILING DATE: 2000-03-01  
 PRIOR APPLICATION NUMBER: 09/823187  
 PRIOR FILING DATE: 2001-03-29  
 PRIOR APPLICATION NUMBER: 60/195792  
 PRIOR FILING DATE: 2000-03-10  
 RESULT 10  
 US-10-453-372-136  
 Sequence 136, Application US/10453372  
 Publication No. US20060003323A1  
 GENERAL INFORMATION:  
 APPLICANT: Alsbrook, et al.  
 TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 FILE REFERENCE: 2140-589\_A  
 CURRENT APPLICATION NUMBER: US/10/453,372  
 CURRENT FILING DATE: 2003-06-03  
 PRIOR APPLICATION NUMBER: 09/789390  
 PRIOR FILING DATE: 2001-02-23  
 PRIOR APPLICATION NUMBER: 60/185967  
 PRIOR FILING DATE: 2000-03-01  
 PRIOR APPLICATION NUMBER: 09/823187  
 PRIOR FILING DATE: 2001-03-29  
 PRIOR APPLICATION NUMBER: 60/195792  
 PRIOR FILING DATE: 2000-03-10

555 QDDNRPFPVITTDNG-----NQLRNFPD-----GAGRH-----

Db 1700 DDD-----VTVLTLNSVEASYTVWQDQVRUNSYQLCNGTLLRVMYANGMGLSFPHSEPHVL 1754

Db 685 -----VSQC-----LKSDGDGKPYTIHQDYDEOGRHTSTSYDL 721

Qy 1755 AGTTPTIGRCMISLPMENGINSIIEWLRKEQIKG-VTIFRKLRHGRNLISI-DYD 1811

Qy 722 TNGR-QOTDPDKVHLMSKSTDNWGO-----IANTHMSY-----GVSEK 759

Db 1812 RNIRTEKLYDDEKFTLRLIVDOVGRPFMLWLPSGLAANVNSYFFNGRLLAGIQRGAMSER 1871

Qy 760 ITVDPIITLTAKQLOQNSNNNNTQGKETTYTPSQQPIQITLDEAGHILQSCTLTTRGWD 819

Qy 1872 TDID-----KQGRIVSRMFADGK-WMSVSYLDSKSNLILQSQROQ-----

Db 1949 PESNASVIFDVSDDRILL-----KYSFLGTGRQVFKYKGKLUSLSEIIVWD-S 1994

Qy 1911 -----IEFDSSDALLAVIMPS---VARH--SMITHTSIGIR---NITYP 1948

Qy 880 TFDGLSLRTQSODGGRMAYTVSAGMDQCPSTVTPQGQFHYQY--QPELDAVLOVAS 937

Db 820 RVRKERTDAIGQCTTYQDYNMVTIQITLPGTIVNRKVAPFESTDTLTDIUNGTSLGQQ 879

Qy 938 NEITQOQSYPNTGAL--LKAVAEQSLPTIYPSPGLKMEINDMKGMSYIWLRLGEN 995

Qy 1995 TAFT----FGYDGTGVLQVNLSOGSGSCTTIRTRKIGL-----VKQIYRFSEGMVN 2045

Qy 996 GYTDLACTIQKISRDTGHRVTOIKDSIKTL---NYDDLN--RHIGS----QVT 1041

Db 2046 ARFDYT-----YHDSNFSRIASIKPVISETPLPDVLRYDESGKVERHFKGKVYIYDIN 2099

Qy 1042 DLATGMLTTVFPD--GLNRBEGRKLUKCDSSHTLDIQLQSMWLTQQLANRTVKNL--GV 1096

Db 2100 QITTAWTLSKHFDTGRKREVYEMPRSLMMWMTQYDSM-----GRVIREKLKGP 2153

Qy 1097 LQRTEQYSYDSRNRLNQYKCDGA-----ECPTDKIGHSTVQNFYDIIGNITACHT 1149

Qy 2154 VANTKTYD-----YDGDQQLQSVAVNDREPTMRY-----SYDNLN----- 2191

Qy 1150 FADGETBDHATKFKAQNPTDPCOLTEVHHTPDMPDNIRLKYDKAGRVINI-----TNHG 1203

Db 2192 -----HLLNPNSVRL-----MP---LRYDJDRTTRGWDQYKIDDG 2227

Qy 1204 -----WTFENPTVDTLGPL-----QNCQGSVYGYDPLNLSQKTOPDLCELYRETMV 1252

Db 2228 YLCQRGSDIFEVNSKGJLITRAYNKASGWSVOTRDGVRRASYYTKN-LGHHLQFYSDH 2286

Qy 1253 NEVR-----NGEMTRL--RTGETIAQQRASKVLTGTDQOS--VILTSKONL 1300

Db 2287 NPTRTRTVYNNHSNEITSLYLVQDGHFAMESSGEEVYVAADNTGPLAVSINGLMK 2346

Qy 1301 QBEASAYGK--HKSTANDASTINGYNGERADPVGVTUHGNGYRSYDPTLMRFPHTD--- 1354

Db 2347 QIQTAYGEIYIDSDNPDFQMVIGFHGGYDPLKLTQVHFTQ--RDYDVLAAGRWMPSDVTW 2404

Qy 1355 -----SLSPFGAGGGINPSYCLGDPINS-DPSGHL---SNQAWTGI----- 1392

Db 2405 KNVGKEPAPF-----NLJMFKSNPPLSELSDLXKVVTDVKSMWLFQFQLSNLIPGFPPRA 2459

Qy 1393 -----GMGTAAGLILITIATGGMATAAGG----IAATA----- 1421

Db 2460 KMWVPPPYELSESQASENGQOLITGVQOTTERHQAQMALEGQVITKKHHSIREKACHW 2519

Qy 1422 -STSTTALAFGL-----SVSDITSVGALEDAKPASSIG----- 1459

Db 2520 FATTPTIGKGIMPAIKEGRVUTGVSSTAS---ED-SKRVASTUNNAYLDKMHYS-EGK 2575

Qy 1460 -----WJSMGMGAQGL-----BSA-----IKGGTKLATHGAFABEDGE 1493

Db 2576 DTHYFVKIGSAQGDLLVLTGTTGRKVLESGVNTVSQPTLINGRTRPFTI---EFQY 2631

Qy 1494 NALKSTSESSRIKVGWT-RSLDRERIVNEEGQVIKOHSGRTDNFMGKGEQAIILVHGDK 1552

RESULT 11  
US-10-453-372-142  
Sequence 142, Application US/10453372  
Publication No. US20060003323A1  
GENERAL INFORMATION:  
APPLICANT: Alisbrook, et al.  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND FILE REFERENCE: 21402-589 A  
CURRENT APPLICATION NUMBER: US/10/453,372  
PRIOR FILING DATE: 2003-06-03  
PRIOR APPLICATION NUMBER: 09/789390  
PRIOR FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: 60/185967  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 09/823187  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 60/199476  
PRIOR FILING DATE: 2000-03-25  
PRIOR APPLICATION NUMBER: 09/863776  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: 60/208263  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 09/939398  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 60/227800  
PRIOR FILING DATE: 2000-08-25  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1609  
SOFTWARE: CurseSeqList version 0.1  
SEQ ID NO 1-42  
LENGTH: 2733  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-453-372-142

Query Match 3.1%; Score 272.5; DB 6; Length 2733;  
Best Local Similarity 19.5%; Pred. No. 3e-08;  
Matches 365; Conservative 251; Mismatches 652; Indels 603; Gaps 99;

Qy 49 GHTIVENG-NIGPTPLTISYSPLNKTDIGEGIGFNFGLSYDKNKLSSLTGENYK-- 104	Db 1066 GHUQKSKQPSQSPNLASTFW--DKTDA-----YQQRVYGLSPAWSY--GPBYETCP 1113
Qy 105 --VIETDKWTKLQOKKLD--NLN--RFEK---DLEKENCYRIIH-KSGDTEVLTG---- 147	Db 1114 SLIWEKRTAIIQGQFELDPSSNLGGWSLDKHLINVKSG--ILHKGTGEMQFLTQQPAIT 1170
Qy 148 -----FNNNAEFLKUP-----KCLLNPNGHAYIDWNEATQPLRNRL-YDDQGDHP 195	Db 1171 TSTMNGRHRISCPSCNGIAEGNKLAPVALYNGIDSP----- 1210
Qy 1114 SLIWEKRTAIIQGQFELDPSSNLGGWSLDKHLINVKSG--ILHKGTGEMQFLTQQPAIT 1170	Db 196 LLNLBYQGKLTILTFPGCGREGTRTELRTLNROL---NSIHNISLGNENTPLTNWSFGYV 251
Qy 1211 -----YVGDFNVIIRRIPSPNNTILER--NKEFKHSNRPANKYLA-VDPVSGSLYVS 1262	Db 252 PIGKNGLQGWITSMTAGCGLKETWYNNQNQGHFPOSANLPVLPYVTLNKQVPCAGQ. 311
Qy 1263 DTNSRRIY---RVKSLSGTKDLAGNSEVAG-----TGBQCLP--DEARCGDGG 1308	Qy 312 AIQAEYTSYTHYUGGSQIWN-----NLKDNLQYGLMTBEYNGSTERRYK---DK 360

Db 1309 AIDA-TLMSPRGIAVDKRNGLMMYFDATMIRKVDQ-NGITSL-LGSMNDTAVRPLSCDS 1364  
 Qy 361 EGHQDQIVRTE-----RTNNYHLLTSECKQOQYQTTETPYYAIGHNFDQSQPQ 412  
 Db 1365 SMDVAQVRLWEPTLAVLNPMNDSLXVL---ENVRIRITEHQVSTA---GRPHC 1415  
 Qy 413 OLPKKTETWRSADNSYRSEBETTFDESGNPLTKVKDKKKOKLISPSTHWEYPAGE 472  
 Db 1416 OVCODISLSKLATHSALESASATAISHCGLVLTETBKKINBLQVTTNGSICLLAGA 1475  
 Qy 473 VDNCPPE-----PYGFTRFVKKIQOTPVSERFKD 503  
 Db 1476 ASPDCKNDVNCCYSGDDAYATDAILNSPSLAVAPDG-TIVIADGNIRRAVSKNPK 1534  
 Qy 504 -ECKPIOTRYSLIGSQSHYTLKLEERHTSATQLNSTL--FQNTDSSELGRUKQEC 558  
 Db 1535 VLNAFNFOYERASRPGEOBLYTVEFADGKIHOTVSLVTGEVLYNFTSTD-NDVTHLI---- 1588  
 Qy 559 TKGENGKTVSVHKFTYTKQDDTQOSH-----SITHDNFTIHSQ----- 600  
 Db 1589 -DINGNSLUKI-----RRDSSGMPRHLMPPDQIITITVGNGLKVUSTONELGLM 1639  
 Qy 601 VSRVYTGRFLSDPTKDVTOMSYDKLGLRLLTRLNSG-----PYANTLTDYELNLT 654  
 Db 1640 TYDGTVGLIATKSDBTGWTFPYDHEGRLTNVRPTGVUTSLHREMEKSTIDIENS 1699  
 Qy 655 QDDNRPPPVITITVNG-----NORNEFD----- 684  
 Db 1700 DDD-----VTVITNLSSVERASYTVQDOVTRNSYOLCNGNTRVNYANGMGISPFSEPHVLT 1754  
 Qy 685 -----VSQC-----LKSDGDKFYTHTQQYDCEOGRHHTSYSDL 721  
 Db 1755 AGTITPTIGRCNTSLPMENGLNSTEWLRKKEQIKGK-VTIFRKLRVGRNLISI-DYD 1811  
 Qy 722 TNGR-QDTDPDKVHLSMSKSYDNGQ-----IANTHWSY-----GVSEK 759  
 Db 1812 RNIKTEKIVYDDHRKETRLTYDQVGRPFMLPLSSGLAAVNVSIPPNGLAGORGAMSER 1871  
 Qy 760 ITVUPITLATAKOLQOSNSNINVPGKEVTRYTSPSOPQIORTFBEAGHQSCHTLDOWD 819  
 Db 1872 TDID-----KOCRIVSRMPADCK-VWSYSYLDKSMVILLOSOROY----- 1910  
 Qy 820 RVRKETDAIGQCTIYQDYNMYNQITUDBGTVWRKAPPSTPLTDIRVNGISLQQ 879  
 Db 1911 -----FEYEDSSDRLLAVTMS---VARH-SMSTHTSGYR---NTNP 1948  
 Qy 880 TFDGLSLRUTQSQDGRGRWVWATYSAGNDOPSTVTPDGDQFHQY--QPELDDAVLOVAS 937  
 Db 1949 PESNASVIRFDYSDGCRIL-----KTSFLGTGRQVFYKKGKUSKSLSEIVD-S 1994  
 Qy 938 NEITQOQSPNPPVTCAL--KAVALGQSLTPVYPSGRLOMENINDMKNSYLMRLGEN 995  
 Db 1995 TAFT--FGYDTEGTVLKMVNQLOSGCPSCTRVRYKIGPL-----VDKOIVRSEEGMVN 2045  
 Qy 996 GYDLGTTGTOKISKDTHGRVTOQKDSKTTL---NYDDLN--RHGS-----QVT 1041  
 Db 2046 AREFT-----YDNEFSRISTKPVISETPPLPVDRYDELSKGVEFRGKFGVIVDIN 2099  
 Qy 1042 DATGHLMLTTVEFD--GJNREICKLKDQSSGHTDIOOSWALKTOQLANRIVKLN--GV 1096  
 Db 2100 QITTAWMLSKHDTHGRIKEVQEMPRSLMMYMTVQDSM-----GRVKGRELKGP 2153  
 Qy 1097 LORTEOYSDRSRNLNOYKCGA-----BCPDKYKHSITVNTONFTYDIYGMNITACHT 1149  
 Db 2154 YANITKTYD-----YGGDQOLQSVAVNDRPTWY-----SYDLNGLM----- 2191  
 Qy 1150 FADCTEDHATPKRANKPTDCQLTQHMPNTRLXKAGRVM-----TDWIG 1203  
 Db 2192 -----HLNAPGNSVRL-----MP---ARYLDRDRITRGLDQVYKIDDG 2227  
 Qy 1204 -----NTENFTYDITGLR-----ONGOSSVYGDPLNRLVSQKTDIUCLEYTRETMV 1252  
 Db 2228 YLCORGSDIFEVNSKGLLTRAYNKASGMSVQYRDGVGRRAVSKTN-LGHHLOYFVSDH 2286

Qy 1253 NEVR-----NGEMTRLL--RTGETTIAQRASKVLTGTDQSOS--VILTSKDQNL 1300  
 Db 2287 NPARITHYHNSNSEITISLYDQGLHFAMESSSGEEYVYASDNTGTLPLAVFSINGLMIK 2346  
 Qy 1301 QRAYSAVK--HISTANDASLIGYNGER&PUSGVTHGNGTYSYDPILMRHPTD--- 1354  
 Db 2347 OQGYTAYGRYIYDSDNPDFOMVGFHGGLYDPLTGLVHPTQ--RDYDVLAGRWTSVDYMW 2404  
 Qy 1355 -----SLSPFGAGGINNPVSYCIGDPINNS-DSBGLH---SMAWTGT----- 1392  
 Db 2405 KNUGKEPRP-----NUJMFKSNPLSSELDKJNQVYVTKSMUWMEGQLSNLPGFFRA 2459  
 Qy 1393 -----GM3JAGLJLTIATGGMATIAAGG-----IAATA----- 1421  
 Db 2460 KMFVUPPYELRSQSASENGQLTGVQOTTERHNQAFALEGOVITKKHASTREKAGHW 2519  
 Qy 1422 -SUSTTAAFGAL-----SUTSDITSVSGALEDAKSASSLG----- 1459  
 Db 2520 FATTTPIGKGMFAIKBGRVITYGSSIAS--ED-SRKVASYVANNAYLDKMHYSIBGK 2575  
 Qy 1460 ---WWSMGMGAAGLA-----BSA-----IKGGTKAATHLGAFABDGE 1493  
 Db 2576 DTWYFKVGSACDSDLVITGTTGRKVLSQSGVNTVSCPTLNGTRBFTN---FPQY 2631  
 Qy 1494 NAIKJSTSBSSRIKWGT-RSLDREIVNEEGOVIKDHSRGYTDNFKGKGBDILVHGDK 1552  
 Db 2632 STULLS-----TRYGLIPDTDEEKAR-----VLDQDORQRALGTAWAKEQOKARDG-R 2678  
 Qy 1553 DGEVYHTSGNK 1563  
 Db 2679 EGSLRWTSGEK 2689

RESULT 12  
 US-10-53-372-145  
 Sequence 146, Application US/10453372  
 Publication No. US20060003323A1  
 GENERAL INFORMATION:  
 APPLICANT: Alsebrook, et al.  
 TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS  
 CURRENT APPLICATION NUMBER: US/10/453,372  
 CURRENT FILING DATE: 2003-05-03  
 FILE REFERENCE: 2140-589 A  
 PRIORITY NUMBER: 09/789390  
 PRIORITY NUMBER: 60185967  
 PRIORITY NUMBER: 2000-03-01  
 PRIORITY NUMBER: 09/823187  
 PRIORITY NUMBER: 2001-03-29  
 PRIORITY NUMBER: 60195792  
 PRIORITY NUMBER: 09/839446  
 PRIORITY NUMBER: 2001-03-19  
 PRIORITY NUMBER: 60199476  
 PRIORITY NUMBER: 2000-03-26  
 PRIORITY NUMBER: 09/863776  
 PRIORITY NUMBER: 2001-05-23  
 PRIORITY NUMBER: 601208263  
 PRIORITY NUMBER: 2000-05-31  
 PRIORITY NUMBER: 09/939398  
 PRIORITY NUMBER: 2001-08-24  
 PRIORITY NUMBER: 601227800  
 PRIORITY NUMBER: 2000-08-25  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 1609  
 SOFTWARE: CurseSequst version 0.1  
 SEQ ID NO 146  
 LENGTH: 2733  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-453-372-146

Query Match Similarity 3.1%; Score 272.5; DB 6; Length 2733;  
 Best Local Similarity 19.5%; Pred. No. 3e-09; Mismatches 652; Indels 603; Gaps 99;

Qy 49 GHIVGNG-NLGPPTPLTYSPLNKNDLGFIGFNGFLSVYDRKNSLISLSTGENYK--- 104  
 Db 1066 GHUOQSFOASPNLASTPIM--DKTDA-----YGQRVGLSDAVVSV-GFVEYRCP 1113  
 Qy 105 -VFTDKTVKLUOKRKL---NL---RPEK---DLKENCYRITH-K-SCDIEVITG---- 147  
 Db 1114 SLIWKEKRTALLLOGFELDFPSNLLGGWSLDKHKHILAKVSG--ILRGKIGENQPLTQQPAII 1170  
 Qy 148 ---FNNAPDLKOP-----IKLLNPAGHAYIDMFEATOPRNLRIYDLDGHIP 195  
 Db 1171 TSIMNGNRKRSTICPSCNGLAEGNKLALPVALAVGIDSL----- 1210  
 Qy 196 LNLNEYQGLKITLTLPGOKEGYTRELAPNLROI---NSIHFSIGNENPLTWSGFT 251  
 Db 1211 -----YVGPFNMYRIPPSRNWTSILEK---NKEPKHSNNPAHKVLA-VDPVGSILVS 1262  
 Qy 252 PIGKNGILGOWWITSWATPGCGKLETKETVNYSNNNGHFFPOSANLPVLPYVIMKQVPGACQ 311  
 Db 1263 DTNSRRYI---RVKSLSGTKDLAGNSEVVAAG----TGEQCLPF---DEARCGDGK 1308  
 Qy 312 AIQEYEVSYSHNVYUGGSINGW-----NKLONLYGMMTEYNGSTESRYK---DK 360  
 Db 1309 AIDA---TLMSPRGIAVDKNGIMYFUDATMTRKVUDQ-NGIISTL-LGSNDLTAVRPLCDS 1364  
 Qy 361 EGHDOIVRBE-----RTYNNHALTSECKQKQNGYIQTETAYAIALIGHNFPDSQSP 412  
 Db 1365 SMDVAQVRLREWPTDLAGVNPNDSLWV-----ENVNLIRLENQVILLA---GRPMIC 1415  
 Qy 413 OLPKKTTETWRSADNSYRSBETETTFDESQNPPLTKVIKOKKTOKLISPSTHWEYYPAGE 472  
 Db 1416 QVRGIDYSLSKLAHSALESSASAISHTQVLYTETDEKKINRLQVTINGECLLAGA 1475  
 Qy 473 VDNCEPPE-----PKFRTRFYKKIQTYPSEKFDP 503  
 Db 1476 ASDCDCNDVNCNCVSGDDAYATDAILNNSPSSLAVAPDG-TIVIADGNIRRASKNPK 1534  
 Qy 504 -EKTIQYRYSLIGQSHVTLKIERHYSATQLINSL---POINTDKSELGRILKQTC 558  
 Db 1535 VLNAFNNOYEASPSGBQBLYVFNADFGIHOYTVLSLVGEVYNTFYSTD-NDVTELJ 1588  
 Qy 559 TKGENGKTYVWHKFTYTKODDTILOQSH-----SITTHDNFTTHRSQ----- 600  
 Db 1589 -DNNNGNSLKI-----RDDSMPRHLMPDNOITLTGTINGGLKVSTQNLIGLM 1639  
 Qy 601 VSRVYTGRLRSDTPDKDVTQMSYKLGRLTRTANSGT-----PYANTLTQYELNL 654  
 Db 1640 TYDGNTGMLATKSDETGWTFYDYHDEGRILNTRPTGVVTSLHREMEKSITIDIENSR 1699  
 Qy 655 QDDNRPFPVITTDING-----NQLENFED-----GAGRH----- 684  
 Db 1700 DDD-----VTVTINLSVEASYTWWODQRNSYOLCNNGTRVAMYANGMGISENSEPRV 1754  
 Qy 685 -VSQC-----LKSDGDSKFYIHTHQYDEQGRHHTSYDYL 721  
 Db 1755 AGTITPTIGRNISLPMENGINSIWEWLKRIKQIKK-VTIFGRKURVGRNLLSI-DVD 1811  
 Qy 722 TNGR-QOTDPKVKHLSMSKDWNNGO-----TANTERSY-----GVSEK 759  
 Db 1812 RNRKETKIDDHKFTIRIYDQVGRPFLWMPSSGLAVAVSYFFNGRLAGLQGAMSER 1871  
 Qy 760 ITVDPIPLTAKQLOGSNSNNWGTGEVTTYPSQDQPIORTFDEAGHLQSCHTLJDGM 819  
 Db 1872 TDID-----KQGRIVSRMPADGK-WWSYSLDKSMWLLIOSQRY----- 1910  
 Qy 820 RVRKETDAIGCTIQYDNTNMRVIQITLPPCTIVRKYAPPSFTDILTDIVNGISLGQQ 879  
 Db 1911 ------IBYDSSPRLAVTMAPS---VARH-SMSTHSIGYR---NIYNP 1948  
 Qy 880 TFDGLSLRTOSSODGGRWAWTYSAGNDQCPSTVITPDQFHYQY--QPBLLDAVLOVAS 937

RESULT 13

; Sequence 150, Application US/10453372  
 ; Publication No. US2006000323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Absorok, et al.  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEAR ACIDS ENCODING SAME, AND METHOD  
 ; FILE REFERENCE: 21402-589 A  
 ; CURRENT APPLICATION NUMBER: US/10/453,372  
 ; CURRENT FILING DATE: 2003-05-03  
 ; PRIOR APPLICATION NUMBER: 09/789390  
 ; PRIOR FILING DATE: 2001-02-23  
 ; PRIOR APPLICATION NUMBER: 60/185967

Db 1949 PESNASVIFDSDGRIL-----KTSPLGTRQFVYKIGKLSKSLRIVD-S 1994  
 Qy 938 NEFTQQFSYNPVGAL-LKAYAEGQSUTPTVPSGKUMENINDMKNSYVWTLGEN 995  
 Db 1995 TAVT-RQYDETGTGVLKVNLOSGGFCTTTRKIGP-----VQOIYIPSEEQGN 2045  
 Qy 996 GYDTLGTIQKLSRDTHGRVTQIKDSSIKTIL---NYDDLN--RHGS-----QVT 1041  
 Db 2046 ARPDY-----YHDNSPRIASIKPVSETPLPVPLVYDREBLSGVFVFGKGSVYIDIN 2099  
 Qy 1042 DLATGHMJLTVERD--GLNREIGRKQKCDGDSGHTLDIQQSWLKTQOLANRIVLN--GV 1096  
 Qy 1097 LQTEQKSYDSRSLRNLQXKCDA-----EPTDKYHSIVTONFTDIYENITACHTT 1149  
 Db 2100 QITTAVMTLSKHFDTORIKEVQYEMERSLWMTVQYDSM-----GRVKGRELKGP 2153  
 Qy 1097 LQTEQKSYDSRSLRNLQXKCDA-----EPTDKYHSIVTONFTDIYENITACHTT 1149  
 Db 2154 YANTKTYD-----YDGDGQLOQSVAVNDPFTWRY-----SDLNGNL----- 2191  
 Qy 1150 PADCTEDRATPKFANPTPCQLTTEVHHTHPDMNDRILKQYDAGRVINI-----TDMHG 1203  
 Db 2192 -----HILNPGNSVRL-----MP-----LRYDARDRITPLGQVYKIDDG 2227  
 Qy 1204 -----NTNFTYDITGRL-----QNGQGSVGYDPLARLVSQRTDIDCELYTRETIV 1252  
 Db 2228 YLCORGSLIFEVNSKGULTRAVNKAQSGVSVQTRYDGVQRRASXTN-LGHHQYFYSLH 2286  
 Qy 1253 NEV-----NGEMHULL--RTGETTIAQORASKVULGTDQSOS--VILTSOKNU 1300  
 Db 2287 NPTRITHYHNHSNSBTSITSLYYDQGHFAMESSGSEBYVVASINTGTPLAVESINGLMK 2346  
 Qy 1301 QEAYSAYK--HKSTANDASILGYNGERADPVSGVTHGNGYRSYDPTLMRFHTPD--- 1354  
 Db 2347 OLOXTAYCERIYDSDNPDFQMVIFGHGGLYDPLTKLWHTQ--RDYDVLAGWNTSPDYNW 2404  
 Qy 1355 -----SLPPFGAGGINPNSYCLGDPINNS-DSFGHL---SMAWTG----- 1392  
 Db 2405 KNYCKEPAFP-----NLYFKSNPLSSELDIKNYVYDVKSMVUMFGQLSNIPGFFRA 2459  
 Qy 1393 ----------GMGJAGLTLTATGGMATIAAGG-----TAATA----- 1421  
 Db 2460 KMVYVPPPVPELSQSASINGQLINGQITGVQOTTERHQAFMALEGQVITKGLHASIREKAGHW 2519  
 Qy 1422 -STSSTALAFGL-----SVSDTSITVSGALEDASPQKASSIG----- 1459  
 Db 2520 PATTPPIGKGIMPAAKQGRVITGVSSIAS--ED-SRKVASVLANAYLQDKMHYSIGK 2575  
 Qy 1460 -----WWSMGMGAGLA-----ESA-----IKGKTQLAHGAPAEGE 1493  
 Db 2576 DTHFVKLGSADGDLVTIGTGRKVLLSGVNVTQSPTLVLNGRTRFTN-----EPQY 2631  
 Qy 1494 NALIKSTSBSSRIKVGWT-RSLOREIVNNEEGQVKIDISRGYTDNFMGKGEQATLNVHDK 1552  
 Db 2632 STILIS-----IRYGLPDTUBEKAR-----VLDQARQALGTAWAKEQQKARDG-R 2678  
 Qy 1553 DGFLYHTBGNK 1563  
 Db 2679 EGSLRWWTBGEK 2689

PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 09/823187  
PRIOR FILING DATE: 2001-03-29  
PRIOR APPLICATION NUMBER: 60/195792  
PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: 09/839446  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 60/199476  
PRIOR FILING DATE: 2000-03-25  
PRIOR APPLICATION NUMBER: 09/863776  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: 60/208263  
PRIOR FILING DATE: 2000-05-31  
PRIORITY APPLICATION NUMBER: 09/939398  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 60/227800  
PRIOR FILING DATE: 2000-08-25  
Remaining Prior Application data removed - See File Wrapper or PAML.  
NUMBER OF SEQ ID NOS: 1609  
SOFTWARE: Curseqlist version 0.1  
SEQ ID NO: 150  
LENGTH: 2733  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-153-372-150

Query Match 3.1%; Score 272.5; DB 6; Length 2733;  
Best Local Similarity 19.5%; Pred. No. 3e-09;  
Matches 365; Conservative 251; Mismatches 652; Indels 603; Gaps 99;

Qy 49 GHIVNG-NLGPPTPLTYSPLAKTDLGFGNFGLSVYDRKNSLISLSTGENY-- 104  
Db 1066 GHURQFSQASPNLASTFW--DXTDA-----YGQRVIGLSDAVUSV-GFEYTCP 1113  
Qy 105 --VIETDKTVKLOCKKL--NL--RFEK---DLKENYRITH-K-SDIEVING---- 147  
Db 1114 SLIWKRTKALLQFELDSINLGGSMSLDKGHLWVSG--ILAKGTGNSQFLQOPATI 1170  
Qy 148 ---FNNAPDLKV-----KJLUNPAGHAYIDWNFRATORPLRNIYDDGDHIP 195  
Db 1171 TSMCNGNRHSISCPSNCNLAEGRKLLAVALAVGIDSLL----- 1210  
Qy 196 LLNTEYQGLKTILTLFPGSREGYTRTELPNROL---NSIRHSLGMENPLTWRSGYT 251  
Db 1211 -----YVGDFNYIRRIFPSRNVTILER--NKGFFKHSNPAHKYLA-VDPVSSLVX 1262  
Qy 252 PIGRNLQWITSMATPGGLKETWVNYSNNQGHFPQSANLPUPIVYLQVPGACQP 311  
Db 1263 DTNSRHY-----RVKSLSGTKDLAGNSTEVAG----TGBCLPF--DEARGDGSK 1308  
Qy 312 AIGAEYSYTSHNYGGSGNGIW-----NKUDNLGYLMTEENYGSTERRVK---DK 360  
Db 1309 AIDA-TLMSPRGIAVDKRGIMYFWDATMIRKQDQ-NGLISLT-LGSNDLTAVRPLSCDS 1364  
Qy 361 EGHQDOIVRLE-----RTTYNTHLTSBCKOONGYQTTERAYAVALIGNNTDSQPSQF 412  
Db 1365 SMDVAQVRLEWPTDVLAVNPMDNSLYLV-----ENNVRIRITEMHNSIA---GRPMIC 1415  
Qy 413 QLPKTKTEWSADNSYRSBITETFDDESGNPLTKVIKDKTOKLISSTHWYYPAGE 472  
Db 1416 QVCPRDYSLSKLAIHSALESASAIAISHGVLYTETBKKNLRQTTNGICLLGA 1475  
Qy 504 --EKPIOTRYSLIGSQSHUTLKIERHYSATOLINSTL--FOINTDSEBLGKLUKQPEC 558  
Db 1535 VLNAPFNQVEAASPQEQELVFNADGIIHQTVSLVGEVLYNFTSTD-NDVTELIT 1588  
Qy 559 TKGEGKTYSVHKEFTYKQDPTIQOSH-----SITTHDNFTHRQ----- 600  
Db 1476 ASOCDCRDVNCNGYSGODAYATDAILSPSSLAWAPDG-TIVIADGNIRAVSKNP 1534  
Qy 473 VDNOPPE----- -PYGFTRFKKLIQTPYDSEKFODP 503  
Db 1422 -STSTLAJFGA-----SVTSRITSIVSGALEDASKASTLIG----- 1459  
Qy 2520 PATTPUICRKIMPAAKSGRVTGCVSSIAS--ED-SRKVAVSLVANAXYLDKOMYSTECK 2575

Qy 601 VRSRTYGRFLSFDTDKVTOOMSTDKGSLULLTRTNSGT-----PYANTLTMYEMLNL 654  
Db 1640 TYDGNTGLLATKSDBTGWTIPYDHEGRLTNVTRPTGVUTSLREMEKSITIDIENS 1699  
Qy 655 QDDRPPFVITITVNG-----NOLRNEFD-----GAGRHI----- 684  
Db 1700 DDD-----VVITVNLSSVEASYTIVQDODVTRNSYQLCNGNTRVWYANGMGSFHSBPHV 1754  
Qy 685 -----VSQCC-----LKDSDSGDKFYTHTQQYDQEGRHHTSYSDYL 721  
Db 1755 AGITIPTIGRCONISLPMENGNSIWERLJKKEQIKGK-VTIPGRKLVRGMLLSI-DVD 1811  
Qy 722 TNGR-QOPTDPDKVHLSMSKSYDWNGQ-----IANTHWSY-----GVSEK 759  
Db 1812 RNIRTEKIYDHRKEPTRLITYDQVGRPFWLWPLSSGLAANVNSYFFNGRLLAGORGAMSER 1871  
Qy 760 ITVPDPILTATKOLOQSNNSINVQKEVTTPSQOPIQITLPDEAGHLOSCHTLIRDGM 819  
Db 1872 TDID-----KQORIVSMPFADCK-VMSYSYLDKSMVLLQSOHQ----- 1910  
Qy 820 VRKETDAIGQCTIYQDYNRNRVQITLDDGTTVNRKYAFTSTDMLTDIRVNGTSLGQ 879  
Db 1911 -----IFEYDSSDRLLAVTMS---VARH--SMSHTSIGIR-----NINP 1948  
Qy 880 TDFGLSLRITQSQDGERVWYATYSAGNDQCPSTVITPDSOFHRYQ-----QPELDDAVQVAS 937  
Db 1949 PESNAVTFDSDPDRIL-----KTSFLGTTGROVQFYKQKUSKLSEIVD-S 1994  
Qy 938 NEITQFSTNPVITGAL-LKAVAEQOSLTPYFSGRLKMNINDMKKOMSYLWTIRGLEN 995  
Db 1995 TAVI--FGYDTEGIVLKMVNLSQSGFCSTTRVKGPL-----VDKQJYRISSEGWN 2045  
Db 996 GYTDLTGTIQKISRDTGRTVQKOSKIKTLL-----NYDDLN--RHIGS-----QVT 1041  
Db 2046 ARFDY-----YHDNSFRIASIKEPVISETPPLPVDLYRDEISGVVERFGKFGVYIYDIN 2099  
Qy 1042 DLATGHEMLTTVEND--GUNREIGRKLCDSGHTLDIQPSWLQKQLANRIVKLN--CV 1096  
Db 2100 QITTAWMTLISKHFDTDGRKEVOYEMFRSLMMWNTVQDSM-----GRVIRKBLKLGP 2153  
Qy 1097 LQRTQYSIDSRNRLNQYKCGDA-----ECPDKYQHSIVTONFTDIYGNITACTH 1149  
Db 2154 YANTTKYTD-----YDGGDQQLQSVAVNDRPTWY-----SYDLNGNL----- 2191  
Qy 1150 PADGETEDHATFKFANPTDPCQLTETVHHTHPDMPNTRKCDKAGRVIN-----TDNHG 1203  
Db 2192 -----HLNPGNSVRL-----MP---LRYDLRDRITRGLDVOYKIDDG 2227  
Qy 1204 -----NTENFTYDTGRL-----ONGOCSVYGDPLNLVSOQKTDLDCBLYYRETMV 1252  
Db 2228 YLCQRGSDIYFENSKGLUTRAYNKASGNSVYRVDGVRASAYCN-LGHHLQFYSULH 2286  
Qy 1253 NEVR-----NEMIRIL--RTGETTIAQORASKVLTGTSQOS--VILTSQDKQNTL 1300  
Db 2287 NPTRITHVHNHSNSEITSLYDLOGLFAMEMSSGEETYVASDNTGTLPLAVFSINGLMK 2346  
Qy 1301 QEASAYGK--HKESTANDASILGYNGERADPVSCHVTHGNGYNSYDPTMLRFHPPD--- 1354  
Db 2347 LOQTTAYGBIVYDPSNPDFOMVIGFHGGLYDPLTKLUVHTPQ--ROYDVLAGRWTSPDYTM 2404  
Qy 1355 -----SLSPFGAGGINPSYCYLDPINRS-DPSCHL-----SWQAWTGI----- 1392  
Db 2405 KNVKEPAP-----NLNFENKSNPLSSEBLDKNVTYDVKSWJWMFGLQFOLNTLPGFEPRA 2459  
Qy 1393 -----GMMGAGLMLITATGCMATAAGG-----IAATA----- 1421  
Db 2460 KMVFVPPYELSESQAENSENGQLITGVQQTTERHNOAFMLEQVITKUHLASIREKAGHW 2519  
Qy 1422 -STSTLAJFGA-----SVTSRITSIVSGALEDASKASTLIG----- 1459  
Db 2520 PATTPUICRKIMPAAKSGRVTGCVSSIAS--ED-SRKVAVSLVANAXYLDKOMYSTECK 2575

RESULT 14  
US-10-453-372-154  
; Sequence 154, Application US/10453372  
; Publication No. US20060003323A1  
; GENERAL INFORMATION:  
APPLICANT: ALBOROCK, et al.  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEAR ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-589 A  
CURRENT APPLICATION NUMBER: US/10/453, 372  
CURRENT FILING DATE: 2003-06-03  
PRIOR APPLICATION NUMBER: 09/789390  
PRIOR FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: 60/185967  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 09/823187  
PRIOR FILING DATE: 2001-03-29  
PRIOR APPLICATION NUMBER: 60/195792  
PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: 09/839446  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 60/199476  
PRIOR FILING DATE: 2000-03-25  
PRIOR APPLICATION NUMBER: 09/863776  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: 60/208263  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 09/939398  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 60/227800  
PRIOR FILING DATE: 2000-08-25  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1609  
SOFTWARE: CuraseqList version 0.1  
SBQ ID NO 154  
LENGTH: 2733  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-453-372-154

Query Match 3.1%; Score 272.5; DB 6; Length 2733;  
Best Local Similarity 19.5%; Pred. No. 3e-09;  
Matches 365; Conservative 251; Mismatches 652; Indels 603; Gaps 99;  
Qy 49 GRIVGNG-NL-QPTPLT-LPSVPLNKDIDFGFQIGENFGLSYDRKNSLSLTSTGENYK-- 104  
Db 1066 QHLFQKSFPQASPNLASTFIW--DKIDA-----YGQRVIGLSDAVSV--GRBYETCP 1113  
Qy 105 --VIEDKTVLQOKKLD--NL--RPEK---DLYCNYTRIHK-SGDIIEFLVQG---- 147  
Db 1114 SLILWERKRTTALQGFELDPSNLGGWSLKHHLNWKSG--ILHGKGTDGENQFLTOPAII 1170  
Qy 148 ---FNNNAFLKVP-----KKLNPAGHAIYDNNEATQFLRNRYDDLDGHDP 195  
Db 1171 TSMENGRRRISSCPSCLNGLEGNKULLAPVALVAGLDSL----- 1210  
Qy 196 LINLEXVOGLIIMLTFPGQEGYRTELRFINQL---NISIHNISLGNEPLTMWSFGYT 251  
Db 1211 ---YVGDFEWYIRRIFPSRNUVISLELR--NKEFKHSNNPAHKTYLA-VPVSGSLIVS 1262  
252 PIGKNGILGQWITSMPAGGKETVWYSNNNQGHFQPQSANLPVLPVYVTLMKQVPGAGQP 311

Db 1253 DTNSRRIY-----RVKSLSGTKDLAGNESEVAG-----TGEOCLPR--DEARCDDCK 1308  
Qy 312 AIQAEYSTSHNVGGGSGNIN-----NKDLNLYGLMTEINYGSTSERRYK---DK 360  
Db 1309 ALD--TMSPRETAVDGLMFVDAWMIRKVDQ-MQIISLT-LGSNDLTAVRPLSDS 1364  
Qy 361 EGDQIVIE-----RTYNYHLLTSECKOONGYQTETAYVATIGHNFDQSQSF 412  
Db 1365 SMDVAQVLEWPFLAVPMNDSLYVL---ENNVIRITERHQVSLIA---GRPHIC 1415  
Qy 413 QLPKKTCTWRSADSASATISHGVLVYITETDDESIGNPLTKVUDKTKYK1ISTSTHWYYPPLGB 472  
Db 1416 QVPGIDYSLSKLATHSALESASATISHGVLVYITETDDESIGNPLTKVUDKTKYK1ISTSTHWYYPPLGB 1475  
Qy 473 VDNCPP-----PVGFTREVKKLTQTPYSEKFDP 503  
Db 1476 ASIDCDKDVNCYCSDAYATDAILNSPSSLAVERP-----TIXADGNTIRAVSKRK 1534  
Qy 504 --EKFIQYRLSITGSQSVTLKTEERHVSATQLNSTL--FOINTDSELGRILKQTCB 558  
Db 1535 VLNAFNQTEASASQEQLYVFNADGIKYTUSLVTGEVLYNFYSTD-NDVTELI---- 1588  
Qy 559 TKGENGKTVSWHFRPTYRQDPLTQSI-----SITHDNFTIHRSO----- 600  
Db 1589 --DNNGNLSK1-----RDRSSGMPPRILIMPMDNQIITLVGTINGLKVSTONLEGLM 1639  
Qy 601 VRSVYTGFLFSDDTDKDVITQMSVXKULGRLLTINSF-----PYANTLTVDYELNL 654  
Db 1640 TYDNTGJLAKTESDEGWTTPYDHECRNLTNTPTCVTSLARREMNSITIDENSR 1699  
Qy 655 QDDNRPPFVTTDVG-----NOLNREED-----GAGRH---- 684  
Db 1700 DDD-----VVVITLSSVEASYTVWQDDQVNNSYOLCNGTLRYANGMGISFSEPVNL 1754  
Qy 685 -----VSOC-----LKDSDGDKFMTHTOOYDDECORHHTSYSDYL 721  
Db 1755 AGITPTGRCNISLPMENGLNSTEWRJKEQIKGK-WTIFGRKLVRVGRNLSI-DVD 1811  
Qy 722 TNGR-QOPTDPKVHLSMMSKSYDWGQ-----IANTHWSY-----GVSEK 759  
Db 1812 RNTITBEKYDTHDRKPTTRITYDQVGRPRTWLPSGGLAAVNVSFNGRLAGLORGAMBER 1871  
Qy 760 ITDVPITLTAKOLOQNSNINQVQKEVTTYTPSQOPIQITLPDEAGHLOSCHTLIRDGD 819  
Db 1872 TDID-----KGDRIVERMFANGK-VNSVSYDQKSMWILLOSROY----- 1910  
Qy 820 RVRKETDATGQCTIYQDMNVRVQITIPDGTVNKRKAPFSTDTLTDIDRVNGISL3Q0 879  
Db 1911 -----IFBYVSDRDLAVTIPS--VARH--SMSTHSISGVIR---NIVNP 1948  
Qy 880 TFDGLSLRQTSQDGERGRWVATYSAWNDPCPSTVTPDQFHYQ---OPBLDDAVLQVAS 937  
Db 1949 PESNAVSDYDGDGRIL-----KTRFLGTSRQVFKYKGKUSLSETVYD-S 1994  
Qy 938 NEITOQPSINPVTCGAL-LKAVAEQGSLTPVYPSGRQIOMENINDMKNSYWLTLGLEN 995  
Db 1995 TAFT--FGYDETTGTVKWNLUQSGGFSCTIRYKGPL-----VDKOIVRFSEGMVN 2045  
Qy 996 GYTDLGTQOKISRDTHGRVTQIKDSSKIKTL----NYDDLN--RHGS----QVT 1041  
Db 2046 ARFDYT-----YHDNSFRIASIKEPVISTPLPVLYRDEISKGVEHKGFCVYIDDIN 2099  
Qy 1042 DLATCHMLTITVVED-GANREIEKRLCDSSGHTUDIOSWLUKTOOLANRIVKLN--GV 1096  
Db 2100 QITTAWMTISKHFTHGRKEVQEMFSLMWMVQDSM----GRVIREBLKGP 2153  
Qy 1097 LQRTQYSYSDSRNRLNQYKCGDA-----ECPPDKYHSIVTQNTFYDIGNITACHT 1149  
Db 2154 YANTTXYT-----YPCDGQLOQSVAVNDRPTRY-----SYDLGNL---- 2191  
Qy 1150 FADGETEDHATFFKFAINTPDCPQLETEVHHTHDPDNIRLKDKAGEGVINI----TDNGC 1203

Db 2192 -----HLNPGNSVRL-----MP---LRYDLDRITRIGDWOKIDDG 2227 ; LENGTH: 2765  
 Qy 1204 ----NTEPFTYDPLGR-----QNGGSSVGYDPLRSLVKSTPTDCELYTREMLV 1252 ; TYPE: PRT  
 Db 2228 YLCQRGSDIPEVNKGLLTRAYNKAQGSWVQYRDGVERRASYTKN-LGHHQFYSDLH 2286 ; ORGANISM: Homo sapiens  
 Qy 1253 NEVR-----NCEMIRL--RTGETTIAQRASKVLTGTSQS--VLTSDKNLI 1300 ;  
 Db 2287 NPTRITHVNHNSEITSLYVQDQHLMFAMESSGEEYTVAINTGIPLAWSINGLMK 2346 ;  
 Qy 1301 QEAVSAYGK--HKSTANDASILGNGERADPVSGVTHLGNQSYDPLMRTPD--- 1354 ;  
 Db 2347 QLOQTYAYGBIYDSNPDFMVIGHGLYDPLTKVHFTQ--RDYDVLAGRNWSPDTMW 2404 ;  
 Qy 1355 ----SISPPGAGGINPYSYCQDPINTS-DPSGHL----SIQAWTGI----- 1392 ;  
 Db 2405 KNGKEPAPP----NLNMFKSNPNPLSSBLDKNLYVTDVSKWVUMFGFOLSNITPGFRA 2459 ;  
 Qy 1393 -----GMGAGLILITATGGMALAAAG----IAATA---- 1421 ;  
 Db 2460 KMVFVPPPYELSESQASENGQVLTGQVQTERINQAFMALESGQVTKKHAIREKAGHW 2519 ;  
 Qy 1422 -SISTTALAFGL----SVTSDITSIVSGALBDAASPASSTIG----- 1459 ;  
 Db 2520 PATTPIGKGIMPAIKEGRVITOVSSIAS--ED-SRKVASVLANNAYLDKMHYSLEGK 2575 ;  
 Qy 1460 ----WWSMGMGAMGLA-----ESAA-----TGGTKLAATHGAFADGE 1493 ;  
 Db 2576 DTMHFVKIGSAQDGLVLTGTTIGRKVULSGVNNTVSQFTLNLNGRTRFTNTI---BFOY 2631 ;  
 Qy 1494 NALKSTSESSRIKMGVT-RSLDRIVRNEEGQVIKDHSGRYTDNFPMQKGEDAILVHDK 1552 ;  
 Db 2632 STULLS-----TRYGLPPTDLOBEKAR----VLDQARQRALGTAWAKEQOKARDG-R 2678 ;  
 Qy 1553 DGFPLYHTSGNK 1563 ;  
 Db 2679 EGSRILWTGCK 2689 ;

RESULT 15  
 US-10-453-372-116  
 ; Sequence 116, Application US/10453372  
 ; Publication No. US20060003323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alabrook, et al.  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 ; CURRENT FILING DATE: 2003-05-03  
 ; CURRENT FILING NUMBER: US/10453-372  
 ; FILE REFERENCE: 21402-589  
 ; PRIOR APPLICATION NUMBER: 09/789390  
 ; PRIOR FILING DATE: 2001-02-23  
 ; PRIOR APPLICATION NUMBER: 6/0185967  
 ; PRIOR FILING DATE: 2000-03-01  
 ; PRIOR APPLICATION NUMBER: 09/823187  
 ; PRIOR FILING DATE: 2001-03-29  
 ; PRIOR APPLICATION NUMBER: 6/0195792  
 ; PRIOR FILING DATE: 2000-03-10  
 ; PRIOR APPLICATION NUMBER: 09/839446  
 ; PRIOR FILING DATE: 2001-03-19  
 ; PRIOR APPLICATION NUMBER: 6/0199476  
 ; PRIOR FILING DATE: 2000-03-25  
 ; PRIOR APPLICATION NUMBER: 09/863776  
 ; PRIOR FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: 6/0/208263  
 ; PRIOR FILING DATE: 2000-05-31  
 ; PRIOR APPLICATION NUMBER: 09/939498  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: 6/0/227800  
 ; PRIOR FILING DATE: 2000-08-25  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1609  
 ; SOFTWARE: Curseqlist version 0.1  
 ; SEQ ID NO 116

---

Qy ; LENGTH: 2765  
 Db ; TYPE: PRT  
 Qy ; ORGANISM: Homo sapiens  
 Db ;  
 Qy ; Best Local Similarity 19.5%; Pred. No. 36-09; Mismatches 652; Indels 603; Gaps 99;  
 Db ; Matches 365; Conservative 251; Mismatches 652; Indels 603; Gaps 99;  
 Qy ; Query Match 3.1%; Score 272.5; DB 6; Length 2765;  
 Db ;  
 Qy ;  
 Db ; 1038 GHIVGNG-NLGPQLPLTISYSPANKTDIGFGIGFNFGISVYDKRNSLISLSTGENYK--- 104 ;  
 Qy ; 105 --VIETDTCVKGQOKKD--NL---RFEK---DLKENCYRITH-KGDIEVLTG---- 147 ;  
 Db ; 1146 SLIWEKRTIALLQGFIELDPSNLGGWSLKHHLINKSG--IHKGTEGENQFLQQPAII 1202 ;  
 Qy ; 148 ---FNNNAFDLKVP-----KLUNPAGEAIYIWNEFEATOPRINRIVDLDGIDP 195 ;  
 Db ;  
 Qy ; 1203 TSIMNGNRRSISCPSCNGLABCNKLAPVALVAGLDSL----- 1242 ;  
 Qy ; 196 LNLBYQOLIKITLLEFGQKESGIRTEFLANQL---NSIHFNSIGENELTWSTGFT 251 ;  
 Db ;  
 Qy ; 1243 -----YUDDFNYIRRIFRSNTSILER--NKEKHNINPHKYIA--VDVSGSYMS 1294 ;  
 Db ;  
 Qy ; 252 PIGENGITGQWITSMTAAPGGLKETVYNNNOCHHPOSANLPLVPLVYTMKQVPGAGQP 311 ;  
 Db ;  
 Qy ; 1295 DTNSRRIX-----RVKSLSGTKOLAGNESEVAG----TGEQCLPF--DEARCQDGK 1340 ;  
 Qy ;  
 Db ; 312 ATOAEYSTISHTHNVGGSENGINN-----NKUDNLYGMLTEYNGSTESRYK---DK 360 ;  
 Qy ;  
 Db ; 1341 AIDA-TUMSPRGIAVDRGKLMYFVDAWMKRDQ-NGILISTU-LGSNDLTAVRPLSCDS 1396 ;  
 Qy ;  
 Db ; 361 EGDQIWR-----RTYNNHLLTSECKOONGYQTETAYAIIIGHNFDQSQQF 412 ;  
 Qy ;  
 Db ; 1397 SMDQAVQLEPLMTAVRPMDSLXV-----ENNVLIRHENHQVSLIA---GRPHC 1447 ;  
 Qy ;  
 Db ; 413 QLRKTKTPTWRSADNSYSEETETTDESIGNPLTKVTKDKTQKTSPTSTHWYYPAGE 472 ;  
 Qy ;  
 Db ; 1448 QVGDIDYLSLSKLAHSALSASAISIASHIGVLUYITERBKKLNRLQWTINGBICLLGA 1507 ;  
 Qy ;  
 Db ; 473 VDQCPPR-----PYGFTRFUKKIOTPYUSEKFIDP 503 ;  
 Qy ;  
 Db ; 1508 ASDCDCDKNDVNCCNCYSGDDAYATDATLNSPSSLAVAPDG-TYIADLGNTRIRAVSKNKP 1566 ;  
 Qy ;  
 Db ; 504 --EKFIQIORYSLIGSQSIFTKIBERHISATOLNSTI--FOYNIDKSESLGRLKQTEC 558 ;  
 Qy ;  
 Db ; 1557 VIANAFNOQEAAAPGEOEYVFNADHGQH-----SITPHDNFTIHRQ----- 600 ;  
 Qy ;  
 Db ; 1621 --DNNGNISKI-----RDRSSGMPRHLMPDNQITLTVGTINGGKUVNSTONLGLM 1671 ;  
 Qy ;  
 Db ; 601 VSRVYTGFLPSDTPDKTIVQSYDKGLJRTIASGT-----PANTYDYNL 654 ;  
 Qy ;  
 Db ; 1672 TYDENTGILAKTESDETGWTFYDHSGRLLNTVTPGCVTLSHREMKSITDIENSNR 1731 ;  
 Qy ;  
 Db ; 655 QDDNRPPPVITTDVNG-----NOJNIED-----GAGR----- 684 ;  
 Qy ;  
 Db ; 1732 DDD-----VVVTLNLSSEASYTWWQDVRNSYVOLCNGTURWYANGMGLSFHSEPHNL 1786 ;  
 Qy ;  
 Db ; 685 -----VSQC-----LKSDSDGDKFYTQHOOYDQSGGRHTSTSVDYL 721 ;  
 Qy ;  
 Db ; 1787 AGITPTIGRCNISLPMENGANSIEWLRKSEQIKGK-VTIFGRKLVRHGRNLISI--DYD 1843 ;  
 Qy ;  
 Db ; 722 TNGR-QCDPDKVHLSMSKSYNNGO-----IANTHNSY-----GYSEK 759 ;  
 Qy ;  
 Db ; 1844 RNRTEKYYDDHHRKFTTLYDQVGRPLWPSGGLAVUNVSYFFNGRLQGAMSER 1903 ;  
 Qy ;  
 Db ; 760 ITWDPIITATKOLOSHSNNTVQGKETTIVTFSQOPIQIUTDEAGHQLSCHLTRLGWWD 819 ;  
 Qy ;  
 Db ; 1904 TDID-----KOGTIVSRMFADGK-VWSYSTLDSKSNVLLQSQDX----- 1942 ;

OY	820 RVRKETDAIQGCTIYQDNYNRVQITLPGDTIVNRYKAPPSIDLITDRNGISLGQ	879
Db	1943 -----1PEYDSDRLLAUTMS-----VARH--SMTHTSIGIR----NTVNP	1980
OY	880 TFDGLSLRQTQSQDSGRWVAYTYSAGNDQCPSTVTPDGOPIHOTY---OPEDDAVLQVAS	937
Db	1981 PESNAVSVIFDYSDDORIL-----KTSFLGTGRQVFYKGKUSKLSSEIVD-S	2026
OY	938 NEITQFESTNPPTVCAI--KAVALSGQLTPYVSSGRIMENTINDMKNSYLMWLGLEN	995
Db	2027 TAFT----FGDTETGVLKWMNLQSCGFSCTRYRKIGPL-----VDKOIVRFSEEGMVN	2077
OY	996 GYTDITGIIQKISKDTHGRVQIKOSSIKTL----NVDLN--RHGS-----QVT	1041
Db	2078 ARFDYT-----YDNDSFRIASIRPVISTPPLPVLYRDEISCKVEHFGKFGVIVYDIN	2131
OY	1042 DLATCHMLTTVEFD--GLNREIGRKLCOSGSHHTDIOQSWLKTQOLANRIVKLN--GV	1096
Db	2132 QITTAWMULSKHDFTHGRIKEVQEMFRSLMMYMTVQIDSM-----GRVIREKLKG	2185
OY	1097 LORTEQSYDSRNINQYKCGA-----BCPDKYGHISIVQNFTDITYGNITACHIT	1149
Db	2186 YANTTKYTD-----YDGDQLOQSVAVNDRPTWRY-----SYDLNGNL-----	2223
OY	1150 PADGETEDHAITFKPANPTDCQLTETRHTIPDMNDNIRUCDKAGRVINI-----TDNHG	1203
Db	2224 -----HLANGPNVSRL-----MP----LRYDLDRITRILGDVOYKIDDG	2259
OY	1204 ----NTENITYDITGRL-----ONGQSVYGDPLNRLVSQKTDITDCELYRETMV	252
Db	2260 YLCQRGSDIDFEYNQSLRIBAQNKAQGWSVQYRIVGGRASASYKTN-LGHHQFYSDLH	2318
OY	1253 NEVR-----NGEMIRLI--RTGETTIAQQRASKVLLTGTDSOQS--VLTSDKONL	1300
Db	2319 NPTRITHVNMHSNSBTSLYDQGLPFAMESSGBEYVVASDNTGIPILAVFSINGLMK	2378
OY	1301 QEAVSAYGK--HKSTANDASLIGYNGERADPVGCVTHLNGYRSYDPTMRFHPPD---	1354
Db	2379 QLOQTYAYGEYDSNPDFOVMVGFHGLYDPLTKLVHFTQ--RDYDVLAGRWTSPPDYTMW	2436
OY	1355 ----SLSPCGAGGINPYSCVLGDPINRS-DPSGHL-----SWQATGI-----	1392
Db	2437 KNVKGKEPAPP-----NLYMKSNNLSSLDLKNVTDYKSLWLMFGEQLSNIPGFPPA	2491
OY	1393 -----GNGTAGLILTATGGMIAAAAGG-----IAATA-----	1421
Db	2492 KMFVPPPYELSESQAENSENGQQLTGQQTTERHNQAFMALEGQVITKKLHASIREKAGHW	2551
OY	1422 -STSTALARGAL-----SVTSIDTISVSGALEBASPKASING-----	1459
Db	2552 FATTPIIGKIMFAKEGRVITGVSSIAS--ED-SRKVAVSVMNNAYVILDKMYSIEK	2607
OY	1460 ----WUSMGMCAGLA-----ESA-----IKGGTKLAATHGAEDGE	1493
Db	2608 DTHYFVKIGADGDVLTVLGTIGKRVLESWNVNTWSOPTULWNGSTRTRFINI----EFOY	2663
OY	1494 NALLKSTSESSRIKNGVT-BSDLRETVNNEGQVTKDHSSRGYTONFMGKCEQATLVHGDK	1552
Db	2664 STILLIS-----IRGLTPPTDIEEKAR----VLDQARQRALGTAKEQQKARDG-R	2710
OY	1553 DGFLHTEGK 1563	
Db	2711 BGSRWLWTEGEK 2721	

Search completed: January 30, 2006, 09:53:34  
 Job time : 29.3373 secs



Db	511	INTD	514		
<b>RESULT 2</b>					
QY	1077	QQSMKTTQQLANRITKLNGVLRQTQYSYNSRNINQYKDGABPPTDKYGHISVTQNFT	1136		
Db	535	EQTYQRNHLIKERITQGRGTRLKRKAFAYPSNRRLIEYTGEARPDYKGAKTHQTF	594		
QY	1137	VDIVONITAGHTFADSTERHATFEEFANPDPQCQLEVEHITHPDMNRLKRYDAGRVI	1196		
Db	595	YDALGMNTKQTDSSGG-RKTVTY-SADPTQDQKVNHDSDPKRKTLEYDAGRMI	652		
QY	1197	NITDHNGNTENPTVYTLGRQ--NG---OCSVYGDPLNLVSO	1235		
Db	653	R-DRAGRT--LRYDALGRILQQNVNGAGAKGQYAVDVNLTVSQ	692		
<b>RESULT 3</b>					
QY	1077	QQSMKTTQQLANRITKLNGVLRQTQYSYNSRNINQYKDGABPPTDKYGHISVTQNFT	1136		
Db	532920	cell wall-associated protein precursor wapa [similarity] - <i>Bacillus subtilis</i>			
C;Species:	<i>Bacillus subtilis</i>				
C;Accession:	S38242				
C;Date:	31-Dec-1993	#sequence revision	31-Dec-1993 #text_change 09-Jul-2004		
C;Accession:	S32920, E65330, T47101				
R;Forster, S.J.					
Mol. Microbiol.	8,	299-310,	1993		
A;Title:	Molecular analysis of three major wall-associated proteins of <i>Bacillus subtilis</i> .				
A;Reference:	roten.				
A;Accession number:	S32919;	MUID:93302506;	PMID:8316082		
A;Accession:	S32920				
A;Molecule type:	DNA				
A;Residues:	1-2334 <FO>				
A;Cross-references:	UNIPROT:Q07833;	UNIPARC:UPI0000060C90;	GB:L05634;	NID:9304177;	PID:10634
R;Kunst, F.; Ogasawara, N.; Mozer, I.; Albertini, A.M.; Altoni, G.; Azevedo, V.; Be C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C.; Ehrlich, S.D.; Emmerсон, P.T.; Entian, K.D.; Bringston, J.; Fabret, C.; Ferrari, Nature 390, 245-256, 1993					
A;Authors:	Foulier, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gai, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, J.; Koetter, P.; Konigstein, G.; Krogius, S.; Kumano, M.; Kurita, K.; Lapidis, A.; Lordin, A.; Lauber, J.; Lazarevic, V.; Lee, S.; Levine, A.; Liu, H.; Masuda, S.; Matsumoto, Y.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanliffe, A.; Schleicher, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sasaki, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tozato, V.; Uchiyama, T.; Winters, P.; Witwit, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, T.; Yoshikawa, H.; Zumstein, E.; Yamashita, H.; Danchin, A.				
A;Status:	preliminary				
A;Molecule type:	DNA				
A;Residues:	1-528 <THI>				
A;Cross-references:	UNIPROT:Q45949;	UNIPARC:UPI0000060C90;	EMBL:X75356;	NID:9407370;	PID:10634
C;Superfamily:	<i>Coxiella burnetii</i> hypothetical protein				
Query Match	10 3%	Score 912.5;	DB 2;	Length 528;	
Best Local Similarity	40 8%	Pred. No. 6.6e-43;			
Matches	222;	Conservative	72;	Mismatches	203;
Indels	47;	Gaps	17;		
QY	18	NEF-FIQANNFTSAASGGVWPRTEGLYNIQITIGHIVGNGNLGPIPLTISYSPIKNDI-75			
Db	2	NELPYTQATNPISAVQGGVDPRTGIFTVNNMLABELTGNDNLGPDPLFTLWHSHTSNC	61		
QY	76	GFGICGFNFGLSVYDVRKNSLISLSTOENSYKVIETDVKVQKQKLNRLFEKDILKENCYRI	135		
Db	62	GFGICGSVGSIVSYDANKKLUISSEBRIKEDWNNSVYVYRQKINPKFEK--JINGYII	119		
QY	136	IHKSDIEVLTG--FNNNAFDLKVPKGLKEVNLPGAHAYIDWNEFEATOPRLRIVDDLDGH	193		
Db	120	KYKNCKTEVLYNKYQGDNLF--LPQKIFSPGLWPLKLWTENRGQVNLTKIEDAD--	173		
QY	194	IPLMLEYQSLIKLTLPQKQEKSYRTELAPLNQRLNISIHFSGNENLITWSGTYPI	253		
Db	174	-VLCKIDYQPSDWARIPEWFGKTESTPQDFVNTEYLWVINKSRSR--LWMSNHYDD	230		
QY	254	KGNGTIGQWITSMTAPGGLKEVNTNSNNNOCHHFPQSNALPVLPVYTMKQVPGACPAI	313		
Db	231	GAGNIT--ITQVKSPGTLTETVNQAGM--RFPDESGKPALPSVNTIQSPGQCPDI	285		
QY	314	QAEXSYTSHNYVGGSN--GIWNKNLNDLG-LMTEYNYSTESTRYKDKEGHDOIYRTE	370		
Db	286	VKEYEVTVSNLYLGASLGKAWNEDBDNTIVVMDDYTSSTEKLIVDNRB---LVSIS	341		
QY	371	RTYNNHILTSBCKQONGYIOTETAYAIGHNFDQSPOSQFLPKKTEWR-SADNSY	429		
Db	342	RYNNYLLISSETTRONSCEVIVETDYAKGLSFQKPFQFOLPKEEKWREPSKNQC	401		
QY	430	RSEITTTFDESGNPLTKVKKTOKITISPSTHMYYPAGEVNN--CPPPEYGFR	485		
Db	402	RSEITTTFDEPEGNLTIKIED-----GKTE-TYDSKGETRKGVIV-CPPPEYGFR	454		
QY	486	FVKKIQTQVSEF---KODPEKFIQYRSLIGOSHVTIKIERHYSATQMLNSTLFO	541		
Db	455	FVKTQIVTPADESEFAPVQOPTYAWQYPCIAQSSLSYAVILOTOETLCSDDVLL--LLT	510		
QY	542	YNTD	545		

Db 1061 SITGPKNKKVTVSYENDLKVTDGTWTSYDYDSEGRLLVKOYKASANSTEAKPVFTRY-- 1118  
 Qy 215 QKGYRTSLRFLRQLNSHFS-----LGMENPLIWSFSPYTPIGNGILGQMTSM 267  
 Db 1119 QYSGHRLKAIKANAKETTYTYSYDADKKTLLMTOPGKRYQYNEAGNP----ICVID 1172  
 Qy 268 APGGIJKETVN-YSNNN-----QGHHRPQSANLPUPLYVYUJMVKQVPGAG 309  
 Db 1173 DAGGKITTNTKVEGNNVYEDVUPNDVGTCATESYQDKDN----VTSVKDAYGT- 1225  
 Qy 310 QPAIQABEYTSYSHNYVGGSGNSGNIWANKUDNLGYGLMTENYGSPESSRYKDQKSDHQVIRI 369  
 Db 1226 ---BTBVKINNDV-----TKMKDEGANVTDIAYDGLDAVSRDQSCKSSAAV 1271  
 Qy 370 ERTTNYVHLTSCEQQNGYIQTETAYAIGHNFPSQSPQQLPKKTEWRSANSY 429  
 Db 1272 YDKYGNQIQLSOKLSASTNLK-----DGSEFAQOKSGNLTASKD- 1311  
 Qy 430 RSEETETPDESSNPLTKVIKDKKTKOIKISPS---TWBYYPPAGEVDNCPPBYGFR 485  
 Db 1312 RKKS-VIADKG-----VLSGCALEVSQSAGTOHYGASSATQTELPTTYTLG 1365  
 Qy 486 FVKKII---QTPDSEFKDPEKPTQY---RYXLLIGSOSHVTUKIEBRHYSATOLLIS-- 537  
 Db 1366 KIKTDLAKSRAVYENIDLRDKDQRKLIWIHEYSALAGKNDWT---KQKIFTTPANAGK 1421  
 Qy 538 --TLFQNTDKEBLGRU-LKQCTCTKENGKTYSVHKFTVKQDDTLQQS-HSITHD 592  
 Db 1422 AVVMEVHDKDOCKGKAWFBEVOLKEGEVSSSYNPVONSFATSENWNVSCASVDEE 1481  
 Qy 593 NFTIHSQRVSRVYTGRLFSDTDKDVTOMSYDKLGRLLTR-----L 635  
 Db 1482 GFDNDVSKAARTSASQGSVTKOTVULGQSAUNDKPVLTLGMSKASSVKFTDEKOYL 1541  
 Qy 636 NSGPYANTLYDYL--NNLDDNRPFV----- 664  
 Db 1542 QANVYIADGSTGIYNAKEFPGSTQFWNRRAAVWIPKTPKPNVUDISLFOKSAIGTVWFDDI 1601  
 Qy 665 -----TTDVGNGLNRNEFGASRHVSQCLKDSDGDKFYTHQODQEGRHT 714  
 Db 1602 RLIBGSLLTKSTYDSNGNTVKEBEDLGATS--TDXDETGK---KTSFTAEGKET 1654  
 Qy 715 STV---SDYLTNGROQTDPKVHLSMSKSYDNWG-QIANT-----HWSKCVSEKI-- 760  
 Db 1655 YTYDQADQJNTNMITSNGTSILH-----SYDKGEHEVSKTRAGAQDQTYKFERVNGKLVK 1709  
 Qy 761 TWDRIITLATAKQLOQNSNNWVQT---GKRVTTPSQQPIQITLFDAGHLSCHTLRD 816  
 Db 1710 TTDPLGNVLAESYDANDNLTKTISPGNEV-----SISYD 1744  
 Qy 817 GWDRVKRKTDAIIGCCTIVYDNTMVRVIQITLPOCTIVRKYAPFSTDPLTDIRVNGSL 876  
 Db 1745 GTDVRVKSISYNGTCKYIFTYDQCN-----ET-SVNUKEON-----TT 1781  
 Qy 877 GQTPDGJSRLTOSQDGGRVWAWTYSAGWDQCPSTVITPDQTHYQVOPELDAVLOVA 936  
 Db 1782 KGTTFDKNKRLTELDGCSQTWVTPSPDKLKT-----SMW-----G 1821  
 Qy 937 SNEBTQOFSYNNPYVAGALLKAVAGSOSLTPPIYPSGRKLMENTINDKMKSYLTIRGLNG 996  
 Db 1822 DOKETNOFTYN-----KUDQMIEMKDTSSSFSDYDEN- 1854  
 Qy 997 YTDLGTQKTSRDTGRTQKQSSIKRTLYDNLHIGSQVPLATGHHMLTTVEFD 1056  
 Db 1855 -----GNVO-----TPICTNGGGTSFSYDERNLVSSLMGDKNGGDLITESVY- 1898  
 Qy 1057 GLNREIGRKLCDSGSSHTLDIQUOSWLUKTOQLANRIVKUNLQRTQYKSDSRRLNQYC 1116  
 Db 1899 -----DANGRTTIN-----ASGKV-----QVEYGLMOLVK--- 1927  
 Qy 1117 DAGECPDTDQHGHSITVTONITYDIVGNITACTTTFADGETD---HATFKEFANPTDCOLTEV 1174  
 Db 1928 -----ETHDGTVI--EVYDGFONRKTV-TTICKDGSKTKVNASFNIMN----QLTKV 1973

Qy 1175 H-----HHTHP-DMPDNI-----RLKCDKAGRVINITDNG 1203  
 Db 1974 NDESISVDKNGNRTSDGKFTYTWDAEIDLAVTKKGDKPFTYKDEKGRIQKTN-G 2032  
 Qy 1204 INTENFTVTLGLQNGGSSVYCDPLMLVLSKTKTDLCELY-YRETMVNEVRNGMIR 1262  
 Db 2033 KVUNYFVIG-----DSLWVYETDADNWNTKSYGD-----SGQILS 2070  
 Qy 1263 LIRTGETTAQQRASKULLTGDQSQSVLITSQDQNLSKQSBAYSAYGR-HKSTANDA--S 1318  
 Db 2071 YTENGKKVYHNTAHCJTIAISDSTGKIV-----AKYQDAWNSNFTKTEASDEYDN 2122  
 Qy 1319 TLYNGERBADPVGVTGNGYRSYDPTLMRHT--PDSLSPFGAGGINPVSYCLGDPIN 1376  
 Db 2123 RIRYAGIQYDEETGLYLMARY-YEFRNGVFLSLPDPGSQDSDLQNGYAGNNNNPVM 2180  
 Qy 1377 RSUPSGHMSWQATGIGMIGIAGLILITATGGMIAAAG---GIAAIASTTALARGA 1432  
 Db 2181 NVDPDGH-W-WMLVUVNGFA-----AYDGYKAVSKGKGWKGAAWAAASNFGPKTIFKG 2231  
 Qy 1433 LSFTSDITSISSGALEDAKPKASSILGWVSMGM---GAGAIAESATGGTKLATHUGAF 1488  
 Db 2232 ASRAYKEFK-----KAVKITCHTRHGINQSIGRNG-----GRGVNLRAKNA- 2273  
 Qy 1489 AEDGENALLKSTSSESSRITKWGVYRSLORE-B-TVRNTERGQVIKOHSGRGYTDNFMGKGEAI 1546  
 Db 2274 -----VRSPKVVIKOPNGATKIVGSKKATVLNKKGKVITAYG----SS 2312  
 Qy 1547 LVHGDKOCFLYTFEGNK 1563  
 Db 2313 RAKGSKHVFHGKGKNC 2329

RESULT 4  
 S38239 :  
 Hypothetical protein - *Coxiella burnetii*  
 CSpecies: *Coxiella burnetii*  
 CDate: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
 CAccession: S38239  
 R.Thiele, D.; Willems, H.; Haas, M.; Krauss, H.  
 Submitted to the EMBL Data Library, October 1993  
 A.Reference number: S38215  
 A.Accession: S38239  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-336 <THL>  
 A.Cross-references: UNIPROT:Q45946; UNIPARC:UPI0000086171; EMBL:X75356; NID:940730; PID:  
 Query Match 4.8%; Score 429.5; DB 2; Length 336;  
 Best Local Similarity 34.3%; Pred. No. 2e-16; Mismatches 91; Gaps 13;  
 Matches 121; Conservative 48; Mismatches 91; Indels 93; Gaps 13;  
 Db 1348 MRHTPDISLSPCGAGGINPYSCLGDPNRSRPSGHLSSQWQMTGIGMIGAIIQTLTITGG 1407  
 Qy 1 1 MRINCPDUSWSPFGAGGINPYSCLGDPNRSRPSGHLSSQWQMTGIGMIGAIIQTLTITGG 60  
 Db 1408 MATAAGGIAATASTSTALARGALSTSDTSISVSGALLEDAKPKASSILGWVSMGMGA 1467  
 Qy 61 TSIAGAAGISATASASISLNVYGTGAAADIASAGALEANPOASATGWLISLG 120  
 Db 1468 AGLAE--SAIKGKTLATHIGAFAEDG-----BNALIKSTSSRSRISKVGT-RS 1513  
 Db 121 PGAVSGLATAARAGKL--ISGLAKGGKIRSKSQSPVQGISTSLSRQDPPLRKGPPHQ 177  
 Qy 1514 LDEBIRVNE-----GOVKDHRSRGTYPF-----NGK 1541  
 Db 178 LSRVTVADESMRPAQGLNYWVKVSSQKSSLGYQHVGADREFIGYEIREPIEFFRRPSTK 237  
 Qy 1512 GEQALI-----VHGDKOGLPLTEG-----NHEGKCPYTRHPEQWVY 1581  
 Db 238 RDVTLISGPHGRVHGDN---WTSQGLRKPDLIBERAFYIEDVQYKG-----QNGR 285

1582	LKDNNNIVDLTQGG-----DKPVLHLSCYCGKGSQGADMAYKIN-RPVIA	1625	Db
286	VKIVDMAGMTESERFGRHRVNVNQHVTILGYCWGRNDEA---LTYHRNLRPVISY	335	Db
RESULT	5		
AT1489	cell wall-associated protein precursor wapa (B. subtilis) homolog lin0454 [Imported] - 1		
C.Species	Listeria innocua		
C.Date	27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004		
C.Accession	AF1489		
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.; D.; Jones, L.M.; Karst, U.			
A;Authors	Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitouram, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,		
A;Title	Comparative genomic of Listeria species		
A;Reference number	AB1077; MUID:2153279; PMID:11676669		
A;Accession	AF1489		
A;Status	preliminary		
A;Molecule type	DNA		
A;Residues	1-2167 <GLA>		
A;Cross-references	UNIPROT:Q92EKS; UNIPARC:UPI00000CC234; GB:AL592022; PIDN:CAC95686.1; C;Genetics:		
Query	Lin0454		
Query Match	43%; Score 386; DB 2; Length 2167;		
Best Local Similarity	19 4%; Pred. No. 1.3e-12;		
Matches	327; Conservative 212; Mismatches 568; Indels 568; Gaps 67;		
QY	17 NNEFFTOQANNFTPSAVSGGVDPRTGKYY---NIQTFLIGHTI-VGNGNL-----GP 59		
718	NNTTSVQSEPATPSIPEALINKQLGKNDWYMTSPVPKGEVVAATNGNLFHLRHTDFNILEGRGP 777		
QY	60 TLPLTLISPLAKTIGFGIGFNFRPLSVPDRKNSLISLSTGENVFKVIEFDKV-----112		
778	SINVRNTENSQDADGATGFGKWTSL-----EEKLVEEEENG-NITVWEESDDKKIIRHPTKG 831		
113	-----KLOQRKLD-----NRTEKDJKENCIRIHKGDIEV 144		
QY	832 DKYEAAPPGIVSBEITWNADGWLKIEBDKSETRFLVDRKLSEKDKTIGNELETTYEYGKLT 891		
145	LTGFNNNAFDL----KVKKLNPAGHAIVDNNEATOPRLNRYYDDLGHDIPLLNLR 200		
892	LRDASGRTVTLLTYEGBLVKEVLKGPPDRKSYTN-----D 926		
QY	201 YQGLIKITLTFPGKKEGYKTRLERLRLRQINSIHFLSLGMENPLTWFSFGYTPIGKNGILG 260		
927	KQELLSISSTARGKLTRYGVDLG-----LTSIVLPKHTERKPYTTEAEE-----972		
Db	261 QWITSMTAGGKETVNYSNNQG-----HHFPOSANLPVLPVYTLMQVPGA 308		
QY	973 EKLTETIDPVQSKTKTISYDQAQQTILTNEKKKKTIVSYDAGN-----PKGEIVDA 1024		
Db	309 GQPAIQAEYSTHTNVGGSSNGIWNWKNDLNYGLMTENYNGSTSRRYKDEGH-----363		
Db	1025 DGLKLTFTTYTESNN-----LVKEVNPKGQDFTYADGNITKAT 1065		
QY	364 DQIVRIERTINYHILTSECKQONQYIQTETAYAIIIGHNFDSDPSQFOLPKTETWR 423		
Db	1066 DAYGTSVYTYNDNNTVS-----STDTE-----GRKTWVTD 1097		
QY	424 SADNSYRSETSEB-----TTEDEBNPL-----TKVIK 450		
Db	1098 GADAASETLATESQVSSVTODAYGPIRGSGELSSGGNLIQNSREBKGAWSNWLQS 1157		
QY	451 DKK-----TQ-----KUISP---STHMXYPAGEEVNCPPPEPYGFPRFVKK 489		
Db	1158 DAKGSMWFDNTOSAPGAGLGGSQSVKLTSEANSTVKGYSSVTOVRVOPETTYTFSAWIKT 1217		
490	-----IIQTPYDSBFKODPEKEIQYRVSLSIGSQSHVTLKIER-HVSAQOLNSTL 539		
QY			
1218	SGTMNADALLIGKIQDANAKVUTDAGW-----QSNRANSIKKNDMWKQJ-----1254		
Db			
QY			
540	PQINTDKEBLGRIL-----KOTECIKGEGKCTSYVHFKEYTKDT 581		
Db			
1265	-TFTKSKOTROVTOVLYLDNBQAPAHKGKPTIWPNVQFERGSVASSVNPVNNSPENHGT 1323		
Db			
582	LQSHSITHDNFT--IHRSOVSRSRYGFLSFDIDTQDITVOMSYD-----KLGRLTR 653		
1324	LPTGWMRGTNTALTOAQVYDQNSHSGSOSAVYFERRKATSEAYTHIVQDVPUVNOKEAKALI 1383		
Db			
634	TMSGTPYAN-----TITYDYEI-----NNIQ-----DDNRPFPVTTD 668		
Db			
1384	SALKSEDRAKNSVATMSNDISVSMGTVYQDCTTSSQSFQFLGNDWNSARAVWKPK 1443		
Db			
669	-----VINGQNL-RNEFGAGRIVSQCLKDSDGKFYI 701		
Db			
1444	PVKNIKVUMFRMFLGKAWFDDVVRVIEGVLTKNEYDASGT-----1487		
Db			
702	HQOYDEQCRHHSITYSULINGRQQTDPKVHLSMSQSYDWWQIANTHWSKVSEKIT 761		
Db			
1488	-TASYDESGRKISKPTDY--GKNTSETDKKNTLYDADNALITKLANGTSVAK 1543		
Db			
762	VDPTLTATKQLOSNNSNNVQTYGKVEVTTVPSQOPI-----QTLFDEAGHQSCHTLT 814		
Db			
1544	YD-----DNGNTTEKVIASGKTKQKIVYEDVWKITAFTA-----LNRT 1585		
Db			
815	RGDGMDRVKETDAl-----SQTCTQYDNTNRTQITLPGTTIVNKYAFGSTDTLITDITY 871		
Db			
1586	KVEYDAAGNETKALIMPNGRVTTESTYDSDADMIGKIKWNDKLAPEQYDENGNTQKWT-BI 1644		
Db			
872	NGLSLQQPTFGDSRLTSQSDGGRWMAVATGNSAGNDQFCSTVITPDQSFHIVQOPELD-----929		
Db			
1645	NSI-----VTDKTYDADNIRTKVAFERGSDVST-----YKQKPTKONK 1682		
Db			
930	--DAVLQASN-----TQOQSYNPVUTGALLKAVAEGOSLTPYPPSGRLKOMENINDK 982		
Db			
1683	GKTDKVGEVAINHGDYTAKTSYND-----RTRVNDGS 1719		
Db			
983	KMSL-WTRGLENGLYDTGQKISRDTGHRVQIQKOSSIKUTLNYD-LNHHGSQL 1040		
Db			
1720	KNAVFEDFERGGINVYTAGCTAANYTDSTQVYNAATISSASGHOIOLDENYTYDAASNR 1779		
Db			
1041	TDAFHMLTTTVERDGLAREIGKUCPSSGHTDIOQSWLKTOOLANLIVKNGVLT 1100		
Db			
1780	TSIONKQDKETTIVYDAVHQ-----LTKEKLPDTGK-----1811		
Db			
1101	EQYSDSRMRNINQVKYQCDGABCP-----DKY-----GHSIVTON-----FTVDTIGNITACHTFADG 1153		
Db			
1812	AYTDGFGNRTQIASSBKTIDASINDQJUSWNGBALTDANGRT-----SDG 1864		
Db			
1154	TEDHATKFKFANTPUPCQQLTVEHHTHPDMDNTRIKYDQKAGRVINITDNGNTENFTYDYL 1213		
Db			
1865	KYTYT-----WDTGDRPLSSST-KKGESEPT-----1889		
Db			
1214	GRLOONGQGSVYGPQLPLRNUJSQKUDLBCLY-----RETMVNEURGEMR-----JLRTGE 1268		
Db			
1890	-----SYTDDNRNLSTKTDVGDUTTMHYDGSITLVLYETGDKVWVQVYSDN 1940		
Db			
1269	TIAQORASKVLTGTDQDOSVILTSKDN-----LSQBAYSAVKG-----HKSTANDAS 1318		
Db			
1941	VLRLONGQGSVYGPQLPLRNUJSQKUDLBCLY-----RETMVNEURGEMR-----JLRTGE 2000		
Db			
1319	ILGNGERA-----DPVSVTHLGNGVSYDPTLMRFTPDPSLSPFGAGGINPY 1367		
Db			
2001	YAGTYDKEIHOYMLARYVEPOCV-----FTAYD-----FTGDEDPP-----QTMGY 2046		
Db			
1368	SYCLGSPINAKPSDPSHLW-----QAVTGIGMIAGLULITMGGMIAAAGGIAAIAASTS 1424		
Db			
2047	NYANNPVMFEDPDCGNVAVWIAAAGYGAPEGGAEVILITKKW-----GFGKAVVKGA 2101		
Db			
1425	TIALFGAL 1433		
QY			

Db	2102 VLGIGFGKL 2110	Qy	726 ---QQTPDPDKVHLSMSKSYDNWGQIAT-THWSYGIVSE----- 758
Db	1204 AGKIVQDGGQAPRQTSVYDENGRTVKAVTISGYVTEWAVDTAYRGDLVLTGAPBGS 1263	Dp	1264 NAVTDARFRTVERRDYAGTQPAGIDYMTRYA-----FDAADRQKSITAHDRSA 1313
Qy	759 -KTVDPPTLTATKQLOSNNSNNVQTKEVTTYTPSQQPIQITLFDENGHLOGSCHTLTRDG 817	C;Species: Streptomyces coelicolor	C;Accession: T37218 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.	submitted to the EMBL Data Library, September 1998	A;Reference number: Z21615	A;Accession: T37218
A;Status: preliminary; translated from GB/EMBL/DDBJ	A;Molecule type: DNA	A;Residues: 1-2183 <OL>	A;Cross-references: UNIPROT:085885; UNIPARC:UPI00000DAD3; EMBL:AL031514; PTDN:CAA20596.
C;Genetics:	A;Experimental source: strain A3(2)	A;Gene: SCOBDB:SC2H4.02	
Query Match 4.3%; Score 381; DB 2; Length 2183; Best Local Similarity 19.8%; Prod. No. 2.5e-12; Matches 366; Conservative 221; Mismatches 669; Indels 594; Gaps 91; Matches 366	Query Match 4.3%; Score 381; DB 2; Length 2183; Best Local Similarity 19.8%; Prod. No. 2.5e-12; Matches 366; Conservative 221; Mismatches 669; Indels 594; Gaps 91; Matches 366	Db	818 W-----DRYRKETPAIGCTIYQXNDYNNEVQIQLPD 849
Qy 99 TGENYKVUTEDKTVKLUQCK-----LDNURFEKDLCENCYRIRITHKSDIEVLRGP 148	Qy 99 TGENYKVUTEDKTVKLUQCK-----LDNURFEKDLCENCYRIRITHKSDIEVLRGP 148	Dp	1314 WTYTYDLEGRQSVTDPDKGTTEYDRAVKSTDGRGEVILFEYDVLGRKTGMWOSA 1373
Db 620 TGDS-----TDSLWLDEIKHKGKRGTPSLSLDAVKFSHVFMEN--RVDGPSDPI----- 666	Db 620 TGDS-----TDSLWLDEIKHKGKRGTPSLSLDAVKFSHVFMEN--RVDGPSDPI----- 666	Qy	850 GITVNRKVKAPFSTDLL-----TDIYVGISLQ-----QTPDGLSRLTQS----- 891
Qy 149 NNNADFLKVPKPKUNPASHAYIYD-----NFEATOPRLNRYDDIDGHOPL 196	Qy 149 NNNADFLKVPKPKUNPASHAYIYD-----NFEATOPRLNRYDDIDGHOPL 196	Dp	1374 KTCGAN-KLAWSPEDTLAKQDQTAVREGEGERGRAYQVKYRDXPLXKVNELTLBAND 1432
Db 667 -LSPERPLRPTVSETGAQQTIVDYLPADCVAGOTMPKADENTKRCFPVWSPYQOBP 724	Db 667 -LSPERPLRPTVSETGAQQTIVDYLPADCVAGOTMPKADENTKRCFPVWSPYQOBP 724	Qy	892 ---DGGRWAYTYSAG-----NDQCPSTVTPD5GFIHYQXQPELDAVTOVASNEIT 941
Qy 197 UNLEYQGLIKILTFLP-GQKEGRTELRFLNLQNLNSIHNSLG-----NEENPL----- 244	Qy 197 UNLEYQGLIKILTFLP-GQKEGRTELRFLNLQNLNSIHNSLG-----NEENPL----- 244	Db	1433 PLVAAGVPARLAESTGTMILDGVTKQAAAPAVAGLSAETVSYT-----DGLQVLTAKGT 1487
Db 725 LDWFQKVPSVSRITLPRGGSEA-----VOHTYEYSGGAMYNDNDPLTAKER 773	Db 725 LDWFQKVPSVSRITLPRGGSEA-----VOHTYEYSGGAMYNDNDPLTAKER 773	Qy	942 ---QPPSYNPVTGALKKAVASGQSPLAPIYSSGRMLMENTNDMKMSYLWTGLRGLNGY 997
Qy 245 TWNSFGYGTPTGKNGTLGOW-----ITSMTAGG-----LKETVUNYENNNOCHHFFPOSANLPV 296	Qy 245 TWNSFGYGTPTGKNGTLGOW-----ITSMTAGG-----LKETVUNYENNNOCHHFFPOSANLPV 296	Db	1488 TGYLIQQAAYSP-----LGDLRQMTIATDPGAKKVVNLND-----YEAGRRLRRTSY 1534
Db 774 TWS-----QNRGKVTHYTOPSGGSTRAKTUVTVLURGMDGDRVLDGDKTP 820	Db 774 TWS-----QNRGKVTHYTOPSGGSTRAKTUVTVLURGMDGDRVLDGDKTP 820	Qy	998 T-DLTGTT-----OKISRDTGHRVQIKOSSKIKITINVDLNRRHIGSQVTDLATHGMITT 1052
Qy 297 PYVTLMKQVPGASOPA-----QAEYTSYHNTYVGGSNGIWNKLNOMLYGMLTE 346	Qy 297 PYVTLMKQVPGASOPA-----QAEYTSYHNTYVGGSNGIWNKLNOMLYGMLTE 346	Db	1535 VTDVHGPMQLSQKQDAGNTSVDAA-----TG-----G-----GTGKADHQ 1576
Db 821 PDRRRKAEVSGVTAATAVTDSDQLAGFQRESVAVGDKEVSGVTDNPKRT----- 871	Db 821 PDRRRKAEVSGVTAATAVTDSDQLAGFQRESVAVGDKEVSGVTDNPKRT----- 871	Qy	1053 VEFDGLNEBEIGRKL-----CDSSGHTIDIOQSWLKTQOLANRIVKUNGVLQRTQEYQSYD 1106
Qy 347 YNGSTESRRYKDXKEGHDOIYRERTYNNHHLTSECKQONGY-----IQTETAY-YAIG 402	Qy 347 YNGSTESRRYKDXKEGHDOIYRERTYNNHHLTSECKQONGY-----IQTETAY-YAIG 402	Db	1577 FTYDG-HKRISAEWTPTADCSTSRT-----VAGIGGAAYWTSQYD 1619
Db 872 -ATQHESYADTEY--YVRIGAS---HARTRITERLNPTDRVRTRTGYDQMPV 921	Db 872 -ATQHESYADTEY--YVRIGAS---HARTRITERLNPTDRVRTRTGYDQMPV 921	Qy	1107 S-----RNRNQYKCDGASCPD-KXG-----HSI-----V-----QNFWDIXENITACHTT 1149
Qy 403 HNFPSQPSQFOLPKTKTETWRADSNS-----YRBBITETFD-----ESGNPL 445	Qy 403 HNFPSQPSQFOLPKTKTETWRADSNS-----YRBBITETFD-----ESGNPL 445	Db	1620 DSGRSKQTEHMSGDDVTEYBGTARGQPHALSATVGAENASYTDEIGN----- 1672
Db 922 SVEBLGDDSVTGBDKCTRHYARNDAAGTALVSRTRVGRACSVTDSALDLPADARP 980	Db 922 SVEBLGDDSVTGBDKCTRHYARNDAAGTALVSRTRVGRACSVTDSALDLPADARP 980	Qy	1150 PADGETHDATKEFANPTDCOLTEVHHTHPDMDPNIRLYDKAGRVINITDHGNTENFT 1209
Qy 446 TKVIKDKKQKQKILSPSTHE--YVPPAGEVNDGEPPEPYGFTREVKKIQTPTUSEFKD 503	Qy 446 TKVIKDKKQKQKILSPSTHE--YVPPAGEVNDGEPPEPYGFTREVKKIQTPTUSEFKD 503	Db	1673 -----TE-----TRPVGRATQDILDNWAEGRLAGVSEPAAGKET 1707
Db 981 GDVVSASKAT--AYDTTWASATOKPTGDA-RNSGRAXGYGGDORPLMTTAVDY-DTL 1035	Db 981 GDVVSASKAT--AYDTTWASATOKPTGDA-RNSGRAXGYGGDORPLMTTAVDY-DTL 1035	Qy	1210 VDTIGRLQNGQSVYGVINPLNRLVSQKTDLCE-----LYRETMVLNEVR-NSEM----- 1261
Qy 504 EKPIQYRSLIGSOSHVTKIEERHYSATQULNISTLFQNTDKSELGRLLKQECTKG 563	Qy 504 EKPIQYRSLIGSOSHVTKIEERHYSATQULNISTLFQNTDKSELGRLLKQECTKG 563	Db	1708 -----GIAVYDAGGDLIRRPPTGEGTIVLGLTTEBVHLKUSGNGAKALSGA 1756
Db .1036 GRPYOVR-----NTDUDTISKTEYTOPPANGPLSTCW-----N 1069	Db .1036 GRPYOVR-----NTDUDTISKTEYTOPPANGPLSTCW-----N 1069	Qy	1222 RLRLTGETIADQARASKVLTGTDQSOSVILPSDKOMNOEASA-----YGG 1309
Qy 564 GKTYSVVKHFTYTKQDDTLQQSHSITTHDNFTIHSQVRSRYGRLFSDTDTKDVTONS 623	Qy 564 GKTYSVVKHFTYTKQDDTLQQSHSITTHDNFTIHSQVRSRYGRLFSDTDTKDVTONS 623	Db	1757 RTYKAGSAVIARTST---AGVSGTKLTFLAGDHHGTSGLAINADTIAFAKRWSTPFGA 1812
Db 1070 AK-----GHRITTVDFAL-----GADLKVTDAANGKVTESA 1100	Db 1070 AK-----GHRITTVDFAL-----GADLKVTDAANGKVTESA 1100	Qy	1310 HKSTANDA-SIIGYNGERADPVGVTUHNGYRSTYPTPLMRHFTPS-LSPFGAGGINP 1366
Qy 624 YDKLGLRLTRL-NSGTPYANTLTYDYEILNLODDNRRPFVITITDVGNGQARNEFDAG 682	Qy 624 YDKLGLRLTRL-NSGTPYANTLTYDYEILNLODDNRRPFVITITDVGNGQARNEFDAG 682	Db	1813 PRCTASGWPDGRFLGPADATGLTQL--GARQYDPDTGFRFLSVDPPLBPDKPNT 1870
Db 1101 YDSLGRVISVWLPNRSRALGKANTVYVGYS--VKSTSLPWWASSAT-LNG-----DQSG 1150	Db 1101 YDSLGRVISVWLPNRSRALGKANTVYVGYS--VKSTSLPWWASSAT-LNG-----DQSG 1150	Qy	1367 YSVCLGDDINRSDPSGHLSWOAWTGIGNGIAGLILTT-----GMAIAAGGAALASTS 1424
Qy 683 RHYWSQCLQKSDGICKFYTHTQYDQEGRHRTSYSDLTLNTR----- 725	Qy 683 RHYWSQCLQKSDGICKFYTHTQYDQEGRHRTSYSDLTLNTR----- 725	Db	1817 YAYASNSTVNTDSPG-----TSDGIG-GIGAIGAIGGUVGAVIGAVAI-TA 1919
Db 1151 YRTYWEIKDS-----LARTRYOVAAPSQGRVIAQTYDGLGULPVTAQADWIIDDTAP 1203	Db 1151 YRTYWEIKDS-----LARTRYOVAAPSQGRVIAQTYDGLGULPVTAQADWIIDDTAP 1203	Qy	1425 TTIALFGALSVSDITSIVSGALEDASKKASSLIGWISMGMAAGLASAIAKGKJLATH 1484
Qy 2006 WLNGGGVILRERODFRG-----GDAFTSILAQDFTSGLRSKQVGRKGTGKGAFAKEV 2059	Qy 2006 WLNGGGVILRERODFRG-----GDAFTSILAQDFTSGLRSKQVGRKGTGKGAFAKEV 2059	Db	1485 LGAPABDGENALIKS-----TSRS-----SRIKQVTRSLDREVRVNEE----- 1523
Qy 1575 PEQDUDYLRKDNNTVDTGCGDKFVHLILSCYGSQGAKMAYINRPU----- 1623	Qy 1575 PEQDUDYLRKDNNTVDTGCGDKFVHLILSCYGSQGAKMAYINRPU----- 1623	Db	1957 -----NFTKSMWDLPPNPBPSOSLLEMALASMPDWIGS-----DPKANRWTTSRSLSF 2005
Qy 2060 GFQVVD-----EGPBPSPWTFKNSLGRGAKDIAVLTINGNGQADAFLG 2107	Qy 2060 GFQVVD-----EGPBPSPWTFKNSLGRGAKDIAVLTINGNGQADAFLG 2107	Qy	1524 -----GQVTKDHS--RGYTDNFMGKGBQAILVFGDK-DGFLYHTEGKHNKG-KGPYTRHT 1574
Qy 1624 AVSNKPTISQGLARIERRDFPLKSTYHSYDPRKILCERTEKVKPKTFRP 1673	Qy 1624 AVSNKPTISQGLARIERRDFPLKSTYHSYDPRKILCERTEKVKPKTFRP 1673	Qy	2006 WLNQGGLVILRERODFRG-----GDAFTSILAQDFTSGLRSKQVGRKGTGKGAFAKEV 2059

**RESULT 7**

Db 2108 TYSKARISS---INKKEGSTLKFSAWNGSD---WRSAHVTVPRSWNP 2150  
**RIBB core protein with extension [imported] - Escherichia coli (strain O157:H7, substrain E90865**  
**C;Species: Escherichia coli**  
**C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004**  
**C;Accession: E90865**  
**R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.**  
**Gasabara, N.; Yasunaga, T.; Kubara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.**  
**DNA Res. 8, 11-22, 2001**  
**A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene**  
**A;Reference number: A96229; MUID:21156231; PMID:11258796**  
**A;Accession: E90866**  
**A;Status: Preliminary**  
**A;Molecule type: DNA**  
**A;Residues: 1-1400 <HAY>**  
**A;Cross-references: UNIPROT:Q8X2F8; UNIPARC:UPI0000D2ADB; GB:BA000007; PIDN:BAB35484.1;**  
**A;Experimental source: strain O157:H7, substrain RIMD 0509952**  
**C;Genetics:**  
**C;Gene: Ecs2061**  
**C;Superfamily: rhaP protein**

Query Match Best local Similarity 4.1%; Score 365; DB 2; Length 1400;  
 Matches 204; Conservative 154; Mismatches 467; Indels 456; Gaps 65;

Qy 418 KTETW-RSADNSYSBEITETFDESGNPLTKV-----KDKKTKITSPSTHNEY 466  
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 218 RLTLYTREAGDLAGETGIVGT-DGAGREFRFLVLTQAQRABEARQHTASLSSPD---- 272  
 Qy 467 YPPAGEVDNCPPERPGPFTRVKKIQTPDFSEKDPEFTIVRYSLGSQSHVTLKIE 526  
 Db -----PREPSDRAPPTDGTBEY-----GPDGIR----- 297  
 527 RHYSATQOLLNSTLFOINTDKSBLGRLLKQTECTIGENGKTV-----VHKETTYKOD 579  
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 298 -LSAVALTHDPAVPESLCAPLARY-----TYTEAGELLAVIDRSNQVRAFTYDQH 349  
 Qy 580 DTLOQSHSITHDNFTIHSQRSRSYTRGLIFSLSDTDKTD1IVTQMAYDKUGRLLTETLNST 639  
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 350 PGRMVVAH-----RYAGR-----PEMYKYYDDIGRVEQVEQLNPAUL 383  
 Qy 640 PYANTLTYDYLNLNODDNEPPFVTTD-WNGNOLRNEDGAGRHVSCLKOSDGSKP 698  
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:  
 384 SY----RYQTB----QDR----ITVTSILNRREVVLHPEGGAG--LKRVVKKLADG-- 425  
 Qy 699 YTIHQIQYDEQGRHHTSYDYLTINGR-QQTDPDKVHLSMSKSYDNWQOINANTHWGV 756  
 Db :|||:|||||:|||||:|||||:|||||:|||||:  
 426 -----SVTHSGGYDAAGRLLAQTD-----AASGRTEVGL 453  
 Qy 757 SEKITVDPITITATQQLQSNNSNNVQGKETYTTPSQQTQITLDEAGHQS---CIT 812  
 Db :|||:|||||:|||||:|||||:|||||:|||||:  
 454 N-VVSGDITDTIDFGREMYFYNDGNQTLAVSPDGLERRAYDEPGLVNSTSRCG 511  
 Qy 813 ITRDGDWRKE----TDAIGQCTIYQVNNTNRYTOITLPGTIVNRYKAPEFSTDLT 867  
 Db :|||:|||||:|||||:|||||:|||||:  
 512 VIRYAXDNPHSELPAATTGSTRONTWRYQGQIAFTDGSQKTRYEDDRFGOMTAH 571  
 Nature 413, 848-852, 2001  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A;Title: Complete genome sequence of a multiple drug resistant *Salmonella* enterica serovar  
 A;Reference number: AB0502; MUID:21534947; PMID:11677608  
 A;Accession: AB0539  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-843 <PAR>

A;Cross-references: UNIPARC:UPI00005A2A4; GB:AL513382; PIDN:CAD08754.1; PID:916501575;  
 A;Gene: STY0324

**RESULT 8**

Db 721 PABOWQYDEBHGMLTEIHLBSCHQAVHYGGYDDKGKLAGEROTVHNPETGELLWIQHET 780  
**RibB-family protein [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi (strain AB0539**  
**C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi**  
**A;Note: this species has also been called *Salmonella typhi***  
**C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002**  
**C;Accession: AB0539**  
**R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaoch, P.**  
**Nature 413, 848-852, 2001**  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A;Title: Complete genome sequence of a multiple drug resistant *Salmonella* enterica serovar  
 A;Reference number: AB0502; MUID:21534947; PMID:11677608  
 A;Accession: AB0539  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-843 <PAR>

A;Cross-references: UNIPARC:UPI00005A2A4; GB:AL513382; PIDN:CAD08754.1; PID:916501575;  
 A;Gene: STY0324

Query Match Best local Similarity 4.0%; Score 353.5; DB 2; Length 843;

Matches 223; conservative 121; Mismatches 308; Indels 327; Gaps 48;

Qy 542 YNTDKSELGR---LLKQFETCKHENGKYSVHKFTVKQDPLQQSHSITHDNFTI 597  
Db 32 HHTDSGEOYRLDMLNLAERSLCLCVTDSMGR 597

Qy 598 RSQVRSRYTGR-----LPSDTIKDVTQMSDKUGRLTRTLNSGTPYANTLTIDE 650  
Db 75 RDEAQMTTFRWSDEERLJLGMTAQGKMRVYVDRGL-----TEHD-P 120

Qy 651 LNNIQDDNRPPF---VITTDVNQNLRNEDF-----GAGHRVSQCL 689  
Db 121 LGRVEQTOWHIPWMIQPETEVDAAGVAMREYDERGNLQAVSDPLHQRTVYGDH-GQW 179

Qy 690 K--PDSGDKSPYTHQTYDQDQG---RH----HTSTYSDYLINGRQDTPQKVHLSMS 738  
Db 180 RITDARGGGDKL---QWNEDGQMLRHTDCSSQTAWYDERTLERVTDAAE-NSTR 232

Qy 739 KSYDNWGQTAINTHWSYGVESEKITUDPVKETDAIGQCTIYQDNYNRVQITLDPGTVNRKX 858  
Db 233 YSYDNGNHTTEVMADGRVYRQD-----ANGRLVYKTSFAG---QI 272

Qy 799 TLFDEBAGHLQSCHTLTDGMDVRVKEKDQNSNNVQTKEVVTTQSQQPQI 798  
Db 273 TRWQ-----RDGGQRVRQRTDATGRRAYEYDAYGRUTLTENGESYRFY- 319

Qy 859 PFSPTLINDIRVNGISLQQTFDQLSLRJOSOD-GGRUWATYSSAGNDQC----- 909  
Db 320 -----DVLDRVIEQTDPBGSRRAYGVNALNAWTAVIYGERG 356

Qy 910 -----STVTPDGQPHYQOPELDAVQV-----ASNEITQF 944  
Db 357 GEIRHGLERDAAGRULAKTIEP---TRTEVRYDAADRLLIRRHHDAEGGEPEVI-RP 412

Qy 945 SYNPVTGALLKAVAEQ-----QSLTPPIYYSGRLKMNINNDMKKMSYLVNLTRGEN 995  
Db 413 SYDSAGNUJSEETAOGLVQHRYDVOGNRKTETQMDGR-----TIRLYY 456

Qy 996 GYTBLTGTTIOKISRDRTHGRVTOIKDSSIKTNTYDDLNRHIGSQVTDATGHMLTTVER 1055  
Db 457 G----SGHLQOI-----NLGRDVISEPTR----- 476

Qy 1056 DGLARBEIGKLCDSQHTDIDQSKLTKTQOLANRIV-XLNGVQRT--EQVSYDSRN 1110  
Db 477 DHLRREVOR---SQR-LDMRRMVDRTGRILTRLKTCMKGMRGVPEFDRREVAYSGODE 531

Qy 1111 LNQYKCDGABCPTDKHGHSITVTONTYDINGNITAC-HTFFADGSTEDHATPKFANPTDC 1169  
Db 532 LLK-----KRHSROGTVDVF-YDTGRITRACRNNEALDSWQYDA--ANLDRR 577

Qy 1170 Q-----LTVWHTHPDMPDPNIRLCKDAGRVIITDNGNTENFTYPTL 1213  
Db 578 QGETAQAQAGCASSVPNFRITSYRGHL-----VRYDEYGRVVERGRGRNG-TQHYRWDAE 628

Qy 1214 GRU-----QNGQSVVG-YDPLNRLVSOQTDLDCLYRETMVNVRENGMIRLR 1265  
Db 629 HRLTEVAVTRGGTVVRYGYDAPRERRVEKGHELDEKGPKPNRTRFLWDMGRLAGECRLR 688

Qy 1266 TGETTIAQGR-----ASKVLTGTD-SQOSVILTSKDKNLSQA-YSA 1307  
Db 689 SSSLIVYSORGSHSERLARVDRRAPEADEATLYYHFDUNGAPEEMTDGGGNIVWEAGYQW 748

Qy 1308 GK--HKSTANDASITGNGYNGERADRVSGVTHLGNGYRSYDPTLWRHTPDLSLSPFG-AGC 1363  
Db 749 GNLTHEKETRPVOONLRFQOGYLDRETGHLH-NLYRFYDPPDICKFISGD--PILAGG 803

Qy 1364 INPSYCYCLGDPBINRSDPG 1382  
Db 804 INLYQYA-PNPLSYIDPLG 821

rhsC protein precursor [similarity] - Escherichia coli (strain K-12)  
C.Species: Escherichia coli.  
C.Date: 12-Sep-1997 #sequence revision 17-Sep-1997 #text\_change 09-Jul-2004  
C;Accession: C64805; 169400; T49812  
R.Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
Science 277, 1455-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A44720; MUID:97426617; PMID:9278503  
A;Accession: C64805  
A>Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1397 <RES>  
A;Cross-references: UNIPARC:UPI00017888F; GB:M29718  
A;Note: the sequence is revised in GenBank entry ECRHSCA, release 114, (PIDN:AAC3073\_1  
R.Oshima, T.; Aiba, H.; Baba, T.; Fujita, K.; Hayashi, K.; Honjo, A.; Iremoto, K.; Inada  
T.; Microbuchi, K.; Mori, H.; Motomura, K.; Nakamura, Y.; Nashimoto, H.; Nishio, Y.; Sai  
DNA Res. 3, 137-155, 1996  
A;Title: A 71.8-kb DNA Sequence of Escherichia coli K-12 Genome Corresponding to the 12.7  
A;Reference number: Z25006; MUID:97061202; PMID:8905232  
A;Accession: T48912  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Residue: 1-1397 <OSH>  
A;Cross-references: UNIPROT:UPI0001338AF; EMBL:D90709; NID:91651305; PIDN:BA35359.1; P  
A;Experimental source: strain K12; Kohara clone 174  
A;Genetics:  
A;Map position: 15.7-16.0  
C;Superfamily: rhsC protein

Query Match 3.9%; Score 342; DB 2; Length 1397;

Best Local Similarity 21.4%; Pred. No. 1.7e-10; Matches 256; Conservative 157; Mismatches 382; Indels 404; Gaps 63;

Qy 418 KTEW-RSADNRSSEITETPDESIGNPKV-----KDKTKQKTS-----PS----- 461

Db 2.18 RTQFPRBAGERSGEINGTCV-DGAGRFLVLTQAQARQQAISGGTSAFPDT 276

Qy 462 -----THMEYYPAGEVDNCPPP-----YGFT-----RIVK 488

Db 2.77 LPGTYBYGRDNGIRLSAWMLTHPEP-----ENLPAPLVRVWTPRGEALAVYDPRNT 331

Qy 489 KIQTPTYSBPFK-----DDEPEKFQYRYSUIGSOSHVTKIEERHYSATOLLNSTL 539

Db 332 QVRSFTYDQKGRGMVAKHTRGRPE--ICYRD---SDGRVTRQINPAGLSYT----- 379

Qy 540 FOXTDKSLSGLRJKQTEC--TKGENGKYSVHKFTTKQDPLQQSHSITHDNFTI 597

Db 380 YQEVENDRITITDSLRREYLTQOCBGG-LKRVVK-----EHADGRT 421

Qy 598 RSQVRSRYTGRLFSDPTDKDVTQMSDKUGRLTRTLNSGTPYANTLTIDEYLANLQDD 657

Db 422 QSQDPA--VGRLQRTDAGRTTEYSPVPTVGLTRIT---TDPGRASAFY----- 468

Qy 658 NRPPFVITTDVNQNLRNEDF-----GDKFYVHTQQYDDEQRH 712

Db 469 NHHSOLTSATGPDSLBRREYDENGRLIQTAPGDTTRYDPSLPCATEDATSR 528

Qy 713 HTSTYSDYLINGRQDTPQKVHLSMSKSYDNWQIAINTHWSYGVESEKITUDVPT-TLTTK 771

Db 529 KTMWSRY--GQNLSTUCSGYVTRYDREFGQVTAIREEGISQYAYDSRQQLINVK 585

Query	Match
QY	3.8%; Score 341.5; DB 2; length 1404;
586 DTKHETRYVE-----YNAAGDLTT--VIADDG-SRNGTQDAGKA	Best Local Similarity 20.2%; Pred. No. 1.8e-10; Matches 255; Conservatory 130; Mismatches 370; Indels 505; Gaps 59;
Db	QY
831 -CTR-----YQDNTNVRVQIOTIPDGTVIUNRKVAPFSIDLTIDKUNGISIGQQTF-	623
C;Species: Escherichia coli	624 ICKTQGLTRSMETDAAGFVIRLTS-----HTTPEVDVJDRLLQTGFGRTOYH 678
C;Accession: E90658	882 -DGILSRQLTOSQDGGRVWAVTYSGNDQCPSTVITPDGQFTHYQOPELDDAVLQVASNEI 940
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and gene	679 HDLTKLKLSEDEGV-----THWHD-EADRHTRTVNGE 714
A;Reference number: A99629; MUID:21156231; PMID:11258796	QY
A;Accession: E90658	941 TQIQSYNPVTGAL--LIKAVAGEQSILTPPIY-YS-PRLKOMENI-----NDM---KKMSY 986
A;Status: preliminary	Db
A;Molecule type: DNA	QY
A;Residues: 1-1404 <HAY>	715 AERWQYDE-RGWLTDISHISEBGRHVTVHYGDSKERLASEHLTVHPQTNELLWQHETRH 773
A;Cross-references: UNIPROT:Q8XED9; UNIPARC:UPI00001653A9; GB:BA000007; PIDN:BAB33660.1;	Db
C;Gentics:	987 LWTGKLENOYTDGTGTOIKISRTHGRITQIKOSSIKITLNLYDNLNRHQSQTDLATG 1046
C;Gene: BCs0237	Db
C;Superfamily: rhsp protein	QY
974 AYNQQLAN--RCIFPSLPAVENLYT-----SGMLGMLKG 808	Db
Db	QY
1047 HMLTTTVEF--DGILREIGKLCDSGGHTDIOQSWLKTQOLANRIVKUNGVIQRTQEQS 1104	376 YHI-LTSEBCKQDQNYIQTETAWYAIIGHNFDSQPSQFOLPKT-KTEWSADNSYSEI 433
809 D-TPLVEYTRDRLHRETIR----SFGRLVLTAAVTPAGQLOSQ-HANSLIS-DRTY 859	245 FHLVLTTOAQRAEVFRKORATLSLSSPAGR--SASSSLVPPTLPACTEYGDNGIRLEA 302
Db	QY
QY	674 803 WLJLHDPAVPDLSLPAAPLARYT-----YTAGELRAVYDPSGTQYRGF 345
Db	QY
1105 YDSRNRLNQXKCDGAEBCPTDKYGHISIVTONFTYIYGNITACHTFAD-----GTEBHA 1158	434 TENTFDSE----GNPLTKVTKDKKTKQKISPSHTHWYYPAGEV--DNCPPPEPGF 483
859 WNDNGELIR----ISSPRO-----TSYSYSTGRLFSDDTDKDIVTOMSYDKLGRILTRTANSGTPYANTLYD 906	303 WLJLHDPAVPDLSLPAAPLARYT-----YTAGELRAVYDPSGTQYRGF 345
Db	QY
1159 TPKPANPTDCQTLTEVHHTPD----MPDN-----IRUKDAGRVINID-----GEDEHA 1200	494 TREVKKLIQTPDSE-----FKDPEKTIQYRSLSIGQSHVTLKTERHYSATL 534
Db	QY
907 GNRFLDP-----ELRFDPSALSMWPNRDIARDAHYLYRGRHGRTEKUDLIPBG 956	589 TWDHDFTHRSOVSRSVYGRGLFSDDTDKDIVTOMSYDKLGRILTRTANSGTPYANTLYD 648
Db	QY
QY	649 VELANLQDNRPRPVITTDUNQNLANEPEGRHVSQCLSDG-----DGKRY 700
Db	QY
1201 ---NGNTENFTYDTLGRILONGQSSVG-----YDPLNRLVSKQHDTLDCEL--- 1244	445 -----DAAGRRTBEYLIMASGAVTAVTGPDR--T 472
Db	QY
957 IRTOBERTHRYHYSQHRLVHYTRTQARPLVSVRVLVPLVPIGRVAKRWRRDLDGM 1016	701 IHTQOYDQEGRHISTSDYLTINGRQQTDPDKVHLSMSKSYDNWQGIAANTHNSYGVSEKI 760
Db	QY
QY	761 VR-YGMSNQROQVSVTYPDGLS-----YDPLNRLVSKQHDTLDCEL--- 1244
Db	QY
1245 -----YT-----RETMVNE-----VRNGEMIRLR---TGTITIAQRA--- 1276	545 YGQILTFDUCSGTYTRYDVRGQOIAVHREGISTYSSYNP-----RG 588
Db	QY
1017 SLSRKRPQTVWYGMDGDRLLTQNDRTRQIYQPSFTPLIRVETATGELAKTORSLAD 1076	521 -DP-----ASELPQGIDATGSTKOM-----AMSR 544
Db	QY
1277 -----SKVLTGTSQQSVLTSKQNLQSEA-----DRVSREMWLASCGLTVAQMOSQ 1303	821 --VRKEDAIQSOCTIVYDNNVRIQTLPOCTIVNRYKAFTSTDUTLTDIVNGISLG 877
Db	QY
Db	878 QQTDFGLSLTTSQDGGERVWAVTYSGNDQCPSTVITPDGQFTHYQOPELDDA--VLO 934
Db	QY
1304 -----YSGAYKEKSTANDAS-----ILGNG 1324	589 Q-----LVSKQDAQGEGETREYSAQDL--TAIVADPGSSEIQLY---DAWGKAVS 634
Db	QY
1134 DPVVTTPARKIHLHYDHRGQLPLJISTEGTAWFAEDEWGNLINEENHQLOQILRPG 1193	935 VAENEITOOFSKMFVPTGALLKAVAEQGSILTPPIYPSGRKLMENINDKCKMSVWTLRGL 994
Db	QY
QY	635 TTQGLLTSRMSGYDPAAGRITVLTNTENGNSCOSTFRDVPDLTEKORGFDGRTORY----- 686
Db	QY
1325 ERADPVSVGTVHLGNGYRSYDPTLFMFTDLSLSPFG-ARGINPYSYCYGDPINRSPDSG 1382	995 NGVYDITGI-----QKISRDTHGRVTOIKSD-- 1022
Db	QY
1194 QQYDPEBESGLY--NHRYYDPLQGXYTQD--P1GLKGQWNFYQQPL-NPISNIDPLG 1246	687 -HYDITGKLTOSBDEGLVTLHYDASDRDITHRTVNGDPAEWQYDEHGWLTLSHTSEG 744
Db	QY
RESULT 10	1023 -IKTLNDDLNKHGSQ--VTDLATGML----- 1049
E90658	Db
RhsG core protein with extension [imported] - Escherichia coli (strain 0157:H7, substrai	745 HRVSVHYGDDKGRLTGRQTVENPETGEMIWEHETGHAYSEGLATRQPDPGLPPVWL 804
C;Species: Escherichia coli	QY
C;Accession: E90658	1050 -----TTTVEF--DGILREIGKLCDSGGHTDIOQSWLKTQOLANRIVL 1093
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and gene	Db
A;Reference number: A99629; MUID:21156231; PMID:11258796	805 TYGSGYLAGMKLGCTPLVYIMRDLRHLRETARSR--GGAYELATAMNTSGQSRHL 861
A;Accession: E90658	QY
A;Status: preliminary	1094 NGVUQRTQPSYVOSRNRLNQYKGDAECPDVKGHISIVTONFTYDIXENITACTHTPDG 1153
A;Molecule type: DNA	862 P---OLDRDYDWNQQL--IRISGP-----ESREVYSDTGRITGVITTAAML 906
A;Residues: 1-1404 <HAY>	1154 TED--HATPKFANP-TDPCQTLTEVHHTPD----MPDNRL-----KYDAGRVIN 1197
A;Cross-references: UNIPROT:Q8XED9; UNIPARC:UPI00001653A9; GB:BA000007; PIDN:BAB33660.1;	Db
C;Gentics:	907 DIDIPIYATDPAGWRLDP-----BLHPSDTIATWPN-RIEADAHVYRIVBYGRILAE 958
C;Gene: BCs0237	Db
C;Superfamily: rhsp protein	1198 ITD-----NNGNTENFTYDTLGRU-----ONGQGSV--YGYPL----- 1229
Db	959 KTRIPEGVIRMEBERTHRYHYSQHRLVFTRIQHGRQEVESRYLYDPLGRRTGKRV 1018



Db	529 KIMMNSRY--GQLISFTDCSGYVTRYDHDIFRGOMTAVRREBEGISQRAYDSRQJOLIAVK 585
Qy	772 QLOQSHSNNVVOTGEKVITYTIPSQQPIQITEDEAMHQ----SCHLTDIGWURKE 824
Db	586 DTQGHETRYE-----YNAIGDLTAVIAPDSRSNRGTYDAWCKAVRT 626
Qy	825 TDAISOCTIVQYDNTNVRIOITLPQGTIVNPKVAPPSTDTLTDTRUNGTSILQOFT-D 882
Db	627 TQG-GLTRSMEDAAAGRVLITSENGS----HTTFRDYLDRJQBTGFDGRTORXHD 680
Qy	883 GLSRUQTOSQDGGRWMAYTISAGNDQCPSTVTPDQFIRHVOYOPELDDAVLOVANNEITO 942
Db	681 LTGKJLURSEDGEGLV-----THWHD-EADRLLTHRVTKGETAE 716
Db	943 QFSNNVITGAL-LKAVEGOSITIYY---PSGL--KHEINIDMKKMSVLM----- 988
Qy	717 RWQYDB-RGMUTDISHISEGRHVA-VHYRYDEKRGTGERQTHPQTEALIWORHTRHA 774
Db	989 -TLRGHENGYDTLGTQIKTSRDTGRTQTKDSSITKTLNDDLRHRIGSQVTDLATGH 1047
Qy	775 YNAQGLAN-RCIPSLPAVENLITG-----SCYLAGMLGD----- 809
Db	1048 MLLTTVEPF--DGLNBRBIRGKLUKDSSGHTLDIQSQWIKTQOLANRIVKLNGVYQRTQOVSY 1105
Qy	810 --TPLYVEYTRDRLRLHETTR-----SFGRYELTTAVTPAGOLQSO-HLNSLIS-DRDYTW 859
Qy	1106 DSRRARNAQKCDGAEPDKYGHSTVTONFITYDYNITIATHTFADGTEPHATEKFANP 1165
Db	860 NDNGELIR-----TSSPRO-----TRSYSTSYTGRLTGVTHTAA-NLDIRIPYA 902
Qy	1166 TDPC--QLETEVHHTHD----MPDN-----IRKUDKGAGRVINID-----NH 1202
Db	903 TDPPGRHLPD-PELHDSTLSMMWDPRDIAHDLYLRYDRGRLTKTDLIPBGVIRTDD 961
Qy	1203 GNTENPYTDLGLONGQ-----GSVVGYDPLNRLYSQKTDTDLCEL----- 1244
Db	962 ERTHRYFDHSOQRHLVHYTRQEEPLVESRVLDPGRVAKWRERDUTGWMISLRK 1021
Qy	1245 ---YY-----RETMVNE-----VRNGEMIRLRLR---TGTTIAQORA----- 1276
Db	1022 PQTWTWCWDGDRLLTTONDRSRRIQTYOPGSFTPLIRVETATGELAKTORNSLADALQOS 1081
Qy	1277 -----SKVLUQTSQSOSVILTSKDQNLISQEA----- 1303
Db	1082 GCHGDGGSVWFPPTVLOMLDRLESEELA--DRVSSESRWLAASCGLTVQEOMKNQMDVYT 1138
Qy	1304 -----YSAYGKHKSTAND--ASILCNGERADP 1329
Db	1139 PARKIHLYCDHGRGLPLALISTREGATAWCAEYDEWGMUNBENPHQDQLQRLPGQDYE 1198
Qy	1330 VSGVTHLGGNYKSYDPTLMRFHTPDLSPLSPFG-AGGINPMSYCLGDPTNRSODSSGHLSWQA 1388
Db	1199 ESGLYY--NRHRHYDPLQGRYITQD--PIGLGGMNLIGYQL-NPISDIDFLGLSMWED 1252
Qy	1389 WTYGIGMGIAGLLITINGMMAAAGGIAAALASTSTALAFGALSVTSIDTISVSGALE 1448
Db	1253 -----AKSG--ACTNCGTUSA-----MGPDKEDSIDSTAY 1283
Qy	1449 DASSPKASSILGWWSMGWAGLAESAIKGKTKATHLGAFADGENALKSTSBSRIKW 1508
Db	1284 DANKINTS-----OSICHEDEKA----- 1301
Qy	1509 GYVTRSLDREIVRNNEQGVIKDHSRGYTDNFMCNGEO-----AILVHGDKDQGFLYHTEG 1561
Db	1302 -----GICKDMSGRVYSTAPRGERKGSYPPNSPCPNGTBKVSAYHTG 1346
Qy	1562 NKGK--GPYTRHTPQHOLVDYKDNII 1587
Db	1347 ADSGEYWDEIFSSGKDEKVK-SRDNNI 1373



B85509 hypothetical protein Z0268 [imported] - Escherichia coli (strain 0157:H7, substrain EDL932) 804  
 C;Species: Escherichia coli  
 C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C;Accession: B85509  
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimaranta, E.; Potamakis, K.; Apodaca, Nature, 409, 529-533, 2001  
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A;Reference number: A85480; MUID:21074935; PMID:11266551  
 A;Accession: B85509  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1404 <STO>  
 A;Cross-references: UNIPROT:Q8XED9; UNIPARC:UPI000000025B; GB:AE005174; NID:912512977; E;C;Genetics:  
 A;Gene: 2028  
 C;Superfamily: rhsF protein  
 Query Match 3.8%; Score 336.5; DB: 2; Length 1404;  
 Best Local Similarity 20.0%; Pred. No. 3.4e-10; Matches 252; Conservative 130; Mismatches 373; Indels 505; Gaps 58;  
 Matches 252;  
 QY 376 YHL-LTSECKQKONGIQTTRAYAALIGHNFDSDQSOPQLKT-KTEWRSADNSYRSHI 433  
 Db 245 FHLVLTQKQRERVERKQRQATLSLSPAGER-SASSSLVFPDTLPAGTEYGANGIRLEA 302  
 QY 434 TETRDDES---GNPLTKVKIKKTKQKTSISPSTHWETYXPAGEV---DNCPEPVGF 483  
 Db 303 WVLTHDPAYFDELPAAPLARYT-----YTAASSELRAVYDVRSGTQVRDF 345  
 Qy 484 TRFVEKTIQIOPYPSE-----FKDDPEKFICRYRSLSIGSQCHVTLKIEERYSATL 534  
 Db 346 A-----YDAEHAGRNVAHRYAGRPS--RYVDDTG-----RVTEL 379  
 QY 535 LN---STLFQINTDKSELGRLLKOTEC-TKGENGKTYSVWHEKPYTKQDDTILQSHSI 588  
 Db 380 VNPGGLDVRPFYRGCGVRTIHDLSNREVLVTEGEGG-LKRVVKK----- 422  
 QY 589 THDNDFNTIHRSSQVRSRVTGFLFSDDTKDITQTMOSYDUGKGRLLRTLNSGTPYANTLTD 648  
 Db 423 -----EHDAGSITRSE-----YDEAGRKKQT----- 444  
 Db 649 YELNNLQDDNRRPPFVITTDVNGNQNLNEFDGASHRHVSQCLKDSDG-----DGKFYT 700  
 Db 445 -----DAAAGRTEVSLHMASGAVTAWTGPDGR-T 472  
 QY 701 IHTQQYDQEGRHHTSYSDVLTNGKQTDKDHVHMSMSYDNGQIANTHWSVGSEKI 760  
 Db 473 VR-YGYNSSQGQTVSYTPGDRHSSRBYDEKGRLAETSRSGE-----TTRYSD---- 520  
 QY 761 TVDPITLTATKQLOQSNNNVQIGKSVTTIPSQQPIQITLDEAHLQSCHTLRDGWR 820  
 Db 521 -----ASEPFTGQDATGSKQM-----AWSR 544  
 Db 821 ---VRKETDAGIQCITYQYDNYNRYQITLDPGTIUNRKAZAPFSTDTLTDIRVNGISLG 877  
 Qy 545 YGQLLTFPTDCSSGGYTRTKEYDHYQQIAVHREGISTYSSNP-----RG 588  
 Db 635 TQQGGLTFRSMWYDAAGRITVLTNENGQSPTFPRDPLRERQRGDRGTRY----- 686  
 Qy 878 QQTFFPLSRSITQSQDQGGRWVWATYSANGNDQCPSTVITPQDFQFHYQYQBLLDDA---VHQ 934  
 Db 589 Q-----LVSKQDAQGRETREYEYSAAGDL-TAIVAPDGSRSEIQY---DAWGKAVS 634  
 Qy 935 WASNETTQQQSYNPVTPGALLKVAEGQSLTPPIYSSGRGLKMINENIDMKKOMSYLWTIRGLE 994  
 Db 687 -HYDLTGKLQSEDEGLVTLWHYDASDRITHRTVNGDPAPQWQDEHGWMLTTSSEG 744  
 Qy 1023 --IKTFLNYDDLNRLRHTGSQ-VTDLATHGM----- 1049  
 C;Comment: the rhs core consist of two distinct parts: a large N-terminal core that is c-

Db 745 HRVSVHYCQYDDKGRLLTGRQTVNPERCEMILWBHETGHAYSBQGLATRQPDPGLLPPVNL 804  
 QY 1050 -----TTTVEF--DGLINNEBIRGKLCDSGSSHTLIQOSMTKQQLANRIVKL 1093  
 Db 805 TCGSGYLQAGMKLQGTPVNEYMDRHTARSP--SGEAVYLTATNTSGOIRSRAHNT 861  
 QY 1094 NGVILQRTQSYDSRNRNLANQYKCDGABCPTDKYGHSTVQNFTYDIVONITACTTFDG 1153  
 Db 852 P---QLDRDYDWNDNGQL-I-RISGPQ-----ESREVYSDIGRLGQHTAANL 906  
 QY 1154 TED--HATPKAMP-TDCOLTEVHHTPDR----MPDNIRL-----KIDKAGRYVN 1197  
 Db 907 DIDPYATDPAGNRLPDP-----ELHPDSTLTAWPNR-RLAEDAHVYVRYDEYGRILAE 958  
 QY 1198 ITD-----NHGNNTNFNTDVLGRJ-----QNGCSV--YGYDPL----- 1229  
 Db 959 KTRIPIPECVTRIMIDERHYTHDSOHNFVHTRIQKBRQPVQSYRVLDPGRRTGKRW 1018  
 QY 1230 -----NRLVSQKTDLDCELYR---ETMLVNUERVGENI 1261  
 Db 1019 RERDLIGMMSLSRKPEEWYGMGDRLTIVQTCRQTRQTYQPGSFITPLRLETFENGQQA 1078  
 QY 1262 RLAR-----TGETMIAO-----QRAKSYLTGTDQSOS----- 1289  
 Db 1079 KARHRSLAEVLQEDTGVLPALAVMLGRLERLROGSVSEBSQQWLAQCLTABQMAQ 1138  
 QY 1290 -----VILISDKQNIQSEQAVASYGK--HSTANDASILGYN 1323  
 Db 1139 LEABYIIRRKLHLHYCDHRLGPLALISFEGETWQGBDEWNLGENSAQHQIQLSILP 1198  
 QY 1324 GERADPVSCTVHGLNGYTSYDPTLMRFHTPDSLSPFG-AGGINPYSVCLGDPINRSDDSG 1382  
 Db 1199 GQYDDEBESGLYNNRNY-YDPLQGRYITQD--PIGEGLMMLYQYPL-NPTEHIDLG 1252

RESULT 15  
 C65159  
 rhsA protein precursor - Escherichia coli (strain K-12)  
 C;Species: Escherichia coli  
 C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
 C;Accession: C65159; S47814; I67402  
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co A;Rose, D.J.; Mau, B.; Shao, Y.  
 Science, 277, 1453-1462, 1997  
 A;Title: The complete genome sequence of Escherichia coli K-12.  
 A;Reference number: A64720; MUID:97426617; PMID:9278503  
 A;Accession: C65159  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-1377 <BLAT>  
 A;Cross-references: UNIPROT:PI6916; UNIPARC:UPI00001338AD; GB:AE000437; GB:U00096; NID:9  
 A;Experimental source: strain K-12, substrain MG1655  
 R;Plunkett, G.  
 submitted to the EMBL Data Library, March 1994  
 A;Reference number: S47666  
 A;Accession: S47814  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-864, 'M' 866-1377 <BLU>  
 A;Cross-references: UNIPARC:UPI000017888B; EMBL:U00039; NID:9666582; PID:AAB18570.1; PI  
 A;Experimental source: strain K-12, substrain MG1655  
 R;Feulner, G.; Gray, J.A.; Kirschmann, J.A.; Lehner, A.F.; Sadovsky, A.B.; Vlazny, D.A.;  
 J. Bacteriol. 172, 446-456, 1990  
 A;Title: Structure of the rbsA locus from Escherichia coli K-12 and comparison of rbsA w  
 A;Reference number: 15-935; MUID:90094253; PMID:2403547  
 A;Accession: 169402  
 A;Status: preliminary; translated from GB:EMBL/DBDJ  
 A;Molecule type: DNA  
 A;Residues: 1-864, 'M' 866-1377 <RES>  
 A;Cross-references: UNIPARC:UPI000017888B; GB:M29716; NID:9147613; PID:9147614  
 A;Experimental source: strain K-12

C:Genetics:  
 A:Gene: rhaA  
 C:Superfamily: rhaP protein  
 C:Keywords: transmembrane protein  
 F1-26/Domain: signal sequence #status predicted <SIG>  
 F1-28/Domain: rhaP protein #status predicted <MAT>  
 P,28-55/Domain: transmembrane #status predicted <TMM>

Query Match 3.7%; Score 311.5; DB 2; Length 1377;  
 Best Local Similarity 20.8%; Pred. No. 6.2e-10;  
 Matches 288; Conservative 170; Mismatches 450; Indels 477; Gaps 75;

Qy 211 LPPQKREGY-RTEARPLNQHQLNSTHNFSLEGENENPLTWGFTPIKGNGLIGQNTSMAP 269  
 Db 126 LPFGEB-DGYSSRSBESLWLR--GSAVLDGHRALWQ----- 160  
 Qy 270 GLGIKETVNYS-----NNNGQ-----HHFPOSANL--PVLPVYTMKOVPGACO- 310  
 Db 161 -ALFELBLRUSPHRYLATNSPOGPWLLLGNCERPEADEVLVAPLPPYRVLTVLGLDFRT 219  
 Qy 311 PAIQAEVSYSTSHNTVGGSNGIWINKLDMLYGLMTYNGSTSRRYKDEGHQIVIR 370  
 Db 220 QTFHREAGCBPSGRITGVTDGAWRH---PRLVL-----TQQR----- 255  
 Qy 371 RTTNNYHLLTSBCKQONGYIQTISTAYAIIIGHNFDQSOPSQFOLPKT---KWTWSADN 427  
 Db 256 -----AERAMQO-----AISG---GTBPSA--PDTPGTYGR--DN 287  
 Qy 428 SYMSBISBTETTPDESGNPLTKVIKRKCTOKIISPSHTHWEVYPAGEVDNCPPPEP---YCFT 484  
 Db 288 GIRLSAVWLTHDPE-----yp-----ENLPAAPLVRCWT 317  
 Qy 485 -----RFVKRIIQTPYSBPFK-----DPBPKFCTYRYSIGSOSHVTKE 525  
 Db 318 PRGBLAVVYDRSKQVRSTFYDDKYGRGRWAHRHGRPB--IIRVYD---SGCRVTEBN 372  
 Qy 526 ERHYSATQOLNSTLRFQWYDKESELGRLLKQTEC---TKGENGKTYSWHKPTYKQDDTIQ 583  
 Db 373 PAGHSYT-----YODERIRITDSJDRREVHTOEGAG--LKRWVKK----- 413  
 Qy 584 QSHSITTHDNFTIHSQVSRSYTRGLFSQDTDKD1IVTOMSYDKLKGRLTRTNSGTPYAN 643  
 Db 414 -----EHADGSVYQSQFDA--VGRRAOTDAAGRTEYSPDVUTGLITRIT--TPDCR 462  
 Qy 644 TLTYYDEUNLNLQDDNRPPFVITDUNGVNOLRNEDGAGRHSQCLKSD----GCKF 698  
 Db 463 ASAIFYNHNRQ-----LTSATGPDSLRLRREYDELGRLOBTAPDGDITRVYDNPB 514  
 Qy 699 YITHTQOYDEQGRHTSTYSDYLTINGRQOTDPDKVHLMSKSYDNWGQIANTHWSGYSE 758  
 Db 515 SDLPCATEDATGSRKTMTWSRY--GQLIISFTOCGSYVYTRIDFRFCQMTAVIRBEGISQ 571  
 Qy 759 KITVDP1-TLTATKQLOQNSNNVNTGKEVTTYTPSQQPTQITPDEAGHQ-----SC 810  
 Db 572 YRAYDSRGCOLIAVQDQHETRVE-----yniAGDLITAVIAPDSR 612  
 Qy 811 HTLTDGMDVRKETDAGTOCQTYQDNYNRVIORTPDGTINRKVAPPSTPLTDIR 870  
 Db 613 NGTQYDAMGKAVRITQOG--GILTRSMYEADAGRVIRLSENGS-----HTTFRVYDRLIQ 666  
 Qy 871 VNGISLGQQT--DGLSLUTQOSDGGGRWAWYIVSAGNDQCPSTVITPDQOFHYQOPEL 928  
 Db 667 ETGFDPGTRTORYHDLTGKLRSBEGLV-----TWHYD-EA 702  
 Qy 929 DDAVLQVASNBTRQQFSNPVPGAL--LKAVALCQSLTPPIYY--PSGRL--KMNENDM 981  
 Db 703 DRUHTHTRVKGTAERWQYDE-RGMLTDISHISSEGHRA--VHVVYDEKGRLTSHRQTVHP 760  
 Qy 982 KQNSYLVW-----TQLSLNGYDLCATIQLKLSRDTHGRVQIKOSSIKTTLYNDDN 1033  
 Db 761 QTEALLWQHTRHAYNAQGLN--RCIPDSLPAVEMLYG-----SGYLAGMKLGD-- 809  
 Qy 1034 RHGSQVTDLATGIMLTTVEP--DGLNREIGRKLCDSQSGHTUDIQQSWLKJQQLANRIV 1091

Db 810 -----TIVLEYVDRDRJRETR-----SFRGVELTVAYTAGQLOQO-- 846  
 Qy 1092 KLNQVLTORTEQYSYDSRSRNLNQYKCDGACRCPDQKXGHISIVTONFTYDIGNITACTHTFA 1151  
 Db 847 HNSLHS-DRDXTWANDNGEFLR-----LSSPRO-----TRYSYSTGRGTVHTA 893  
 Qy 1152 DGTEHDHATPKFAMPDPO--QLTEVHHTHPD---MDN-----IRLKDAGRVIN 1197  
 Db 894 -----NLDIRIVPATDPAQNRLD-PELAHPDSTLSMWPDNRJARDAAHLYRYDRHGRUTE 947  
 Qy 1198 ITD-----NHGNTNTMFVTLGRONGQ-----GSVYGMPLNLKVQKDT 1239  
 Db 948 KTDLIPCGVIRTDBERTHRYHYDOSRHLVHYTRQVEBPLVSEYRLDPLGRVAKRWVR 1007  
 Qy 1240 LDCEI-----YY-----RETMVNE-----VRGEMIRLRL---TGETI 1270  
 Db 1008 REEDLTGMSLSRKQVTVWGMGDRUTIQDRTTQTIQVPGSFPLRIVETAGELA 1067  
 Qy 1271 IAOORA-----SKVLTGTDQOSVIL-----TS 1294  
 Db 1068 KTOKRSLADALOOSGGEDGGSVVFPVVLQMLDRLESBILADRVSSEERRWLASCGLTV 1127  
 Qy 1295 DKON-----LSQA-----YSAYCHKSTAND--AS 1318  
 Db 1128 OMONOMDPVYTPARKIHLYCDHRLGPLALISKEGTTEWCAEYDEWGNLNLBNRPHQOQ 1187  
 Qy 1319 ILGNGERADPVGVTGNGYRSYDPMTLMRHTPDSISPFQ-AGGINPYSCLGDPINR 1377  
 Db 1198 LILRGQYDDESGLYY--NRHRYDPOGRVYTD--PIGKGGMNFIQYL-NPVTN 1241  
 Qy 1378 SDPSG 1382  
 Db 1242 TDPLG 1246

Search completed: January 30, 2006, 09:49:50  
 Job time : 40.174 sec

THIS PAGE LEFT BLANK

Page 1

Run on:	January 30, 2006, 09:40:48	{ Search time 133.373 Seconds (without alignment) } 8849.971 Million cell updates/sec
OM protein - protein search, using sw model		
Title:	US-09-889-874A-23	
Sequence:	1 VYIKFLKLFRRTMSDNNEF . . . . . PRKILGLRTEKVVKPKIFPR 1673	
Scoring table:	BLOSSUM62	
Searched:	Gapop 10.0 , Gapext: 0.5	
Total number of hits satisfying chosen parameters:	2166443	
Minimum DB seq length:	0	
Maximum DB seq length:	200000000	
Post-processing:	Maximum Match 0% Listeing first 45 Summaries	
Database :		
1: UniProt_05_80;*		
2: uniprot_trembl;*		
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.		
Result No.	Score	Query Length DB ID
1	8815	99.3 1660 2 Q9EVRT_XENBV
2	1632	18.4 1625 2 Q7MB38_PHOIL
3	1615	18.2 1578 2 Q4ZT05_PSESY
4	1614	18.2 1562 2 Q883V9_PSESM
5	1598	18.0 1385 2 Q7NWU2_CHRYO
6	1583	17.8 1735 2 Q7NA47_PHOIL
7	1553	17.5 773 2 Q9S6J1_COXBU
8	1530.5	17.2 1774 2 Q52B80_COXBU
9	1490	16.8 1682 2 Q4ZTP7_PSESY
10	1416	15.9 1632 2 Q88LP7_PSEPK
11	1214	13.7 709 2 Q45948_COXBU
12	1037	11.7 1290 2 Q88LP2_PSEPK
13	932.5	10.5 982 2 Q883W6_PSESM
14	920.5	10.4 1669 2 Q87VG6_PSESM
15	919.5	10.4 505 2 Q45905_COXBU
16	919.5	10.4 526 2 Q9X626_COXBU
17	916.5	10.3 526 2 Q52B83_COXBU
18	912.5	10.3 528 2 Q45942_COXBU
19	884	10.0 589 2 Q45902_COXBU
20	648	7.3 762 2 Q93IJ8_9VIBR
21	552	6.2 370 2 Q7NVR2_CHRYO
22	486	5.5 389 2 Q7NA00_CHRVO
23	478.5	5.4 310 2 Q45904_COXBU
24	450.5	5.1 2217 2 Q8TP72_METAC
25	450	5.1 2817 2 Q7NW05_CHRVO
26	439	4.9 2554 2 Q7QRS5_LEPTIC
27	432	4.9 2334 2 WAP_A_BTC3
28	429.5	4.8 336 2 Q45946_COXBU
29	420.5	4.7 2364 2 Q82R58_STRAW
30	418.5	4.7 1835 2 Q6MD84_PARW
31	414.5	4.7 2221 2 Q73B23_BACCI
Result No.	Score	Query Length DB ID
1	8815	99.3 1660 2 Q9EVRT_XENBV
2	1632	18.4 1625 2 Q7MB38_PHOIL
3	1615	18.2 1578 2 Q4ZT05_PSESY
4	1614	18.2 1562 2 Q883V9_PSESM
5	1598	18.0 1385 2 Q7NWU2_CHRYO
6	1583	17.8 1735 2 Q7NA47_PHOIL
7	1553	17.5 773 2 Q9S6J1_COXBU
8	1530.5	17.2 1774 2 Q52B80_COXBU
9	1490	16.8 1682 2 Q4ZTP7_PSESY
10	1416	15.9 1632 2 Q88LP7_PSEPK
11	1214	13.7 709 2 Q45948_COXBU
12	1037	11.7 1290 2 Q88LP2_PSEPK
13	932.5	10.5 982 2 Q883W6_PSESM
14	920.5	10.4 1669 2 Q87VG6_PSESM
15	919.5	10.4 505 2 Q45905_COXBU
16	919.5	10.4 526 2 Q9X626_COXBU
17	916.5	10.3 526 2 Q52B83_COXBU
18	912.5	10.3 528 2 Q45942_COXBU
19	884	10.0 589 2 Q45902_COXBU
20	648	7.3 762 2 Q93IJ8_9VIBR
21	552	6.2 370 2 Q7NVR2_CHRYO
22	486	5.5 389 2 Q7NA00_CHRVO
23	478.5	5.4 310 2 Q45904_COXBU
24	450.5	5.1 2217 2 Q8TP72_METAC
25	450	5.1 2817 2 Q7NW05_CHRVO
26	439	4.9 2554 2 Q7QRS5_LEPTIC
27	432	4.9 2334 2 WAP_A_BTC3
28	429.5	4.8 336 2 Q45946_COXBU
29	420.5	4.7 2364 2 Q82R58_STRAW
30	418.5	4.7 1835 2 Q6MD84_PARW
31	414.5	4.7 2221 2 Q73B23_BACCI
Result No.	Score	Query Length DB ID
1	8815	99.3 1660 2 Q9EVRT_XENBV
2	1632	18.4 1625 2 Q7MB38_PHOIL
3	1615	18.2 1578 2 Q4ZT05_PSESY
4	1614	18.2 1562 2 Q883V9_PSESM
5	1598	18.0 1385 2 Q7NWU2_CHRYO
6	1583	17.8 1735 2 Q7NA47_PHOIL
7	1553	17.5 773 2 Q9S6J1_COXBU
8	1530.5	17.2 1774 2 Q52B80_COXBU
9	1490	16.8 1682 2 Q4ZTP7_PSESY
10	1416	15.9 1632 2 Q88LP7_PSEPK
11	1214	13.7 709 2 Q45948_COXBU
12	1037	11.7 1290 2 Q88LP2_PSEPK
13	932.5	10.5 982 2 Q883W6_PSESM
14	920.5	10.4 1669 2 Q87VG6_PSESM
15	919.5	10.4 505 2 Q45905_COXBU
16	919.5	10.4 526 2 Q9X626_COXBU
17	916.5	10.3 526 2 Q52B83_COXBU
18	912.5	10.3 528 2 Q45942_COXBU
19	884	10.0 589 2 Q45902_COXBU
20	648	7.3 762 2 Q93IJ8_9VIBR
21	552	6.2 370 2 Q7NVR2_CHRYO
22	486	5.5 389 2 Q7NA00_CHRVO
23	478.5	5.4 310 2 Q45904_COXBU
24	450.5	5.1 2217 2 Q8TP72_METAC
25	450	5.1 2817 2 Q7NW05_CHRVO
26	439	4.9 2554 2 Q7QRS5_LEPTIC
27	432	4.9 2334 2 WAP_A_BTC3
28	429.5	4.8 336 2 Q45946_COXBU
29	420.5	4.7 2364 2 Q82R58_STRAW
30	418.5	4.7 1835 2 Q6MD84_PARW
31	414.5	4.7 2221 2 Q73B23_BACCI

Db	361 NYHILITSECKQOQNYIOTETAYAIIIGHNFDQSOPSQFQLPKTKEWRSADNSYSEI	420	QY	1514 LDEBIVRNNEEQVIKDHSRGYTDNFMGKGBQALVHGDKGGLIHTSGNKHKGPYTRH	1573
Qy	434 TETPDESGNPLTKWKVKDKKKTKIISPSTHWEXYPAGEVDNPPEPYGFTREVKKIQT	493	Db	1551 LDEBIVRNNEEQVIKDHSRGYTDNFMGKGEQALVHGDKGGLIHTSGNKHKGPYTRH	1560
Db	421 TETPDESGNPLTKWKVKDKKKTKIISPSTHWEXYPAGEVDNPPEPYGFTREVKKIQT	480	QY	1574 TPBOLVDTLKDDNTVDLJLOGGKPKVHLSCYGSAGDKMAYKINRPIAVSNKPTISQ	1633
Qy	494 PYDSERPKDPEPKTQYRYSILGSQSHVTLKIEERHYSATQULNSTLQYNTDKSBLGJL	553	Db	1561 TPBOLVDTLKDDNTVDLJLOGGKPKVHLSCYGSAGDKMAYKINRPIAVSNKPTISQ	1620
Db	481 PYDSERPKDPEPKTQYRYSILGSQSHVTLKIEERHYSATQULNSTLQYNTDKSBLGJL	540	QY	1634 GLARIERDFPLSTYMSYDPRKLIGRTTEKTKPKTPRP	1673
Qy	554 QKTECTKGENGKTVSVHKEITYKQDITQOSHISITTHDNFTHSQRSRYGRFLSDT	613	Db	1621 GLARIERDFPLSTYMSYDPRKLIGRTTEKTKPKTPRP	1660
Db	541 QKTECTKGENGKTVSVHKEITYKQDITQOSHISITTHDNFTHSQRSRYGRFLSDT	600			
Qy	614 DTKDQIVTOMSYDKUGLRLTRTNSGTYANTLTDYELNLQDDNRPPFVITTDVNGO	673			
Db	601 DTKDQIVTOMSYDKUGLRLTRTNSGTYANTLTDYELNLQDDNRPPFVITTDVNGO	660			
Qy	674 LRNERDAGGRHVSQCLKDSDGDKEYITMQYDQEGRHTSTSVDYLNGRQDPKV	733			
Db	661 LRNERDAGGRHVSQCLKDSDGDKEYITMQYDQEGRHTSTSVDYLNGRQDPKV	720			
Qy	734 HLSMSKSYDNGQIANANTHNSYGVSEKITUDPITLTAKOLQSNENNVTGEVYTTSQ	793			
Db	721 HLSMSKSYDNGQIANANTHNSYGVSEKITUDPITLTAKOLQSNENNVTGEVYTTSQ	780			
Qy	794 QPIOITLFDEAGHLOSCHTLTDGMDVRKETDAIGQCIYQDYNVRVQITLPGTIV	853			
Db	781 QPIOITLFDEAGHLOSCHTLTDGMDVRKETDAIGQCIYQDYNVRVQITLPGTIV	840			
Qy	854 NRKVAPPSTDLLTDIRNGISLQGQFDSLRLTQSOPGGRWAWYTSAGNDQCPSTW	913			
Db	841 NRKVAPPSTDLLTDIRNGISLQGQFDSLRLTQSOPGGRWAWYTSAGNDQCPSTW	900			
Qy	914 TPDGOFIHKOQPELDADDIVQVASHEITQOFSYNVTGALLKAVABQGSPLIPVPSGR	973			
Db	901 TPDGOFIHKOQPELDADDIVQVASHEITQOFSYNVTGALLKAVABQGSPLIPVPSGR	960			
Qy	974 KMENTINDMKQSYLWTLRGLLENGYDLGTTQKLSRDTHGRVTOQKDSIKTNYDDN	1033			
Db	961 KMENTINDMKQSYLWTLRGLLENGYDLGTTQKLSRDTHGRVTOQKDSIKTNYDDN	1020			
Qy	1034 RHISQVTDLATTGMLTTTFEDGUNREGRKLDSQSGHTDOSWLMQQLANRIVKL	1093			
Db	1021 RHISQVTDLATTGMLTTTFEDGUNREGRKLDSQSGHTDOSWLMQQLANRIVKL	1080			
Qy	1094 NGVLORTEQYSYDSRANLNOQKCDGABCPCDQKHSIVTONFTDIGNITACHTPADG	1153			
Db	1081 NGVLORTEQYSYDSRANLNOQKCDGABCPCDQKHSIVTONFTDIGNITACHTPADG	1140			
Qy	1154 TEDHATPKFAMPPTDQCOLTEVHTHDPDMQDNLRLKYDKDAGRVINNTDNGNTENPTYDTL	1213			
Db	1141 TEDHATPKFAMPPTDQCOLTEVHTHDPDMQDNLRLKYDKDAGRVINNTDNGNTENPTYDTL	1200			
Qy	1214 GRLQNGQGSTGYDPLNRLVSKQPTPLDCSLYREMLVNEVRNGEMIRLRTERTIAQ	1273			
Db	1201 GRLQNGQGSTGYDPLNRLVSKQPTPLDCSLYREMLVNEVRNGEMIRLRTERTIAQ	1260			
Qy	1274 ORASKULLGTDSQSVILTSQDKONIQSSEA SAYGKHKSTANDASILGNGERADPVGV	1333			
Db	1261 ORASKULLGTDSQSVILTSQDKONIQSSEA SAYGKHKSTANDASILGNGERADPVGV	1220			
Qy	1334 THLGNGYRSYPTLMRHTPDSLSPRGAGGINPVSCLGDPINRDPGSHLSWOAWTGIG	1393			
Db	1321 THLGNGYRSYPTLMRHTPDSLSPRGAGGINPVSCLGDPINRDPGSHLSWOAWTGIG	1380			
Qy	1394 MGIAGLILTTAGMIAAAGGIAAIASTSTTAAFGALSIVTSVGALESQKLA	1453			
Db	1381 MGIAGLILTTAGMIAAAGGIAAIASTSTTAAFGALSIVTSVGALESQKLA	1440			
Qy	1454 ASSLGWSMKGAGAELAESIIGGGKLA	1513			
Db	1441 ASSLGWSMKGAGAELAESIIGGGKLA	1500			

RESULT 2  
 OTMB38 PHOLL ID OTMB38; PHOLL PRELIMINARY; PRT; 1625 AA.  
 DT 01-MAR-2004 (TREMBREL, 26. Created)  
 DT 01-MAR-2004 (TREMBREL, 26. Last sequence update)  
 DE similar to the nematocidal protein 2. Probable membrane protein.  
 GN OrderidocusName=plu222;  
 OS Photobacterium luminescens (subsp. laumontii).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Photobacteriales;  
 OX NCBI\_TAXID=141679;  
 RN [1]  
 RP R  
 RC STRAIN=T701;  
 RX MEDLINE-22957627; PubMed=14528314; DOI=10.1038/nbt886;  
 RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,  
 RA Taucourt S., Bocs S., Bourboulle-Eude C., Chandler M., Charles J.-F.,  
 RA Dasse B., Derouze R., Deruelle S., Freyssinet G., Gaudriault S.,  
 RA Medigue C., Lanos A., Powell K., Siguer P., Vincent R., Wingate V.,  
 RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;  
 RT "the genome sequence of the entomopathogenic bacterium Photobacterium  
 luminescens".;  
 RL Nat. Biotechnol. 21:1307-1313 (2003).  
 DR PMID: BX571866; CABE14515.1; -; Genomic\_DNA.  
 DR Photobacter sp. plu222; -;  
 DR InterPro; IPR006330; YD; PFam; PF05593; RHS repeat; 7.  
 DR TIGRFAMS; TIGR01643; YD\_repeat\_2x; 3.  
 KW Complete proteome.  
 SQ SEQUENCE 1625 AA; 1801B MW; BFA62F4BCB3EA97 CRC64;

Query Match 18.4%; Score 1632; DB 2; Length 1625;  
 Best Local Similarity 31.0%; Pred. No. 1.4e-79; Mismatches 653; Indels 204; Gaps 59;

Matches 494; Conservative 245; Mismatches 653; Indels 204; Gaps 59;

QY 1B NEFTIQANQNTSAVSGGGVPRTGLYNTICLIGHLVGN--GNLQPTLPITLSPNPLKTDI 75  
 DB 4 NDIVSNAPNGSYINTGVPRTGQYSANINNITIRPNVNQIQQV-LNLSPSLTIVLN 61

QY 76 GPGIGFNGFSLSVDRKNSLSSLGENK--VIETDKVLUQOKKLNRLPEKDLCNG 132  
 DB 62 GFGIGWRSFLTMPLDVKTLFSRNGEOKCKPKLPPNNNIDISKDKKLDRVYK-LDNT 120

QY 133 YRITHSGDIEVLTGFFNNNAFDLKVKPKLNPAKHATIDWNFFATQPLRNLYDDGH 192  
 DB 121 FVVKNGKIGETURKRIGSS-DIAKTVALEFPDG-----EVLDLINS 161

QY 193 DIPLNLEYQGLIKITLITL-FPSQKEGYRTELREL-NROLNISHNFSLGNENPFTWSRGY 250  
 DB 162 RFAUSEKIRMTGKYLKUNYSGN--CTSVEPDDNNISAKAFDYRNDLITVTPVY 218

QY 251 TPIGK-----NGLGOMITSMAPGALKETMYNNMNGHIFPQSANLPVLP 297  
 DB 219 DASGPIDSARFKMVTQTLKGIF-FVISAFTPCTGYVLELYSYKEN--GHKVDTSEIYAA 275

QY 298 YVTLMKQVQAGAQAEIQAQSYTS-HNYVGGSNGI-WNKLDNLGYLMTEYNGESTR 355

Db 276 ALTIQ--PGNGQPAISKSYESVSVHNFGLGGSSRTSFDSQQNLYLVGKYYTSSR- 330  
 Qy 356 RYKDEKGHDQIVIERTVNNHILTSECKQONGYIQTETAYAIIIGNFDSQPSQFCLP 415  
 Db 331 -RVTANGONVISVTERVFDFKFLHTKEAQTDQNRKRITABTMDPSHSFSBEPNTQOP 388  
 Qy 416 K-----TKTEWTWTSADNSYRSESETTEFDDESSNPLTKVIKDKKIQKUTSPSTHWEYPP 469  
 Db 399 SHVLTTRYDQLOT---NWSREBNVIKSDDWMGNTLL-ITETSGIQK---EVVYKV 436  
 Qy 470 AGEVNDCCPBPYGSPTRFVKCIIOPYDSEBFKDPPEKFQYRKLIGSOSHVTLKIEHHY 529  
 Db 437 NGEGNNCPADLGPSPRLFKSVTKGSPDAAQSVAANRVTSYTYOKLPTGAVYKK-BEVs 494  
 Qy 530 SATOLINSLT--PQY--NTDSKSELGRULKQECTKGENGKYSVWHKFTVKQDPLQ 583  
 Db 495 KASBTIDSKVIRTVNVSNSPNTKSH-GSLAKITSVMNNQ---OTVTPPKVYSDENT 548  
 Qy 584 OSHAITTHDNPTIHSQRSRSYTORLFSDFDTDKIVTOMSYDKUGRLJRTLSGTPTVAN 643  
 Db 549 TNSWTGEGTGHMSEKMSVYTHQLRKUDWVHITDOSGRIGQIDPGTKEI 608  
 Qy 644 TLTYDYLNLQDNPPEPVTTDVGNOQLRNEFDGAGRHVSQCLKSDG-----D 695  
 Db 609 KRSTIVQPGGENDFWP-WMIEDSOGIRKTHYDGMRCSIEBQDDGWGWTSGYQ 667  
 Qy 696 GKPVTTHQYQDECGRHMSTSYLTLN---GRQQTPDKHLSMSKSYDWGQIANT 750  
 Db 668 GTVRKVLRQDVQGLVKEISNDWLWDISANPLTRLTP-LVTKTYQDNGNRYST 725  
 Qy 751 HWSYGVSEKITVDPILTAK-----QLONSNNNVQTKETVVTYTPSQOPQITLDEA 804  
 Db 726 EYGDGRBLEIHPITRITQVKGMLNITQDNP-----EOPASIKVYPP 773  
 Qy 805 GHQSCHTLTRDGMDVRKETDATGOCTIYODVNRYQITDGTUVNRYKAPFSTD 864  
 Db 774 GATISRTTRYDGFRTWETDABGIAQTOIEYDLPDRVKKLUDRTLESAYFSHE 833  
 Qy 865 LMDIRVNGISLSQTFDGLSRLTOSPDGGRWVATYASGNOCPSVTWIP-DGQFTHY 922  
 Db 834 LISALNVNGTQLCASLVLVDPBGIQGRVTRDVGKTEVLYGSQDK-PGSVTPAINKONDY 892  
 Qy 923 QYQBLDDAVLQVASNETTQOFSPNVTGALLKAVAGESSLTT-YFSGRKLHENID 980  
 Db 893 LY--ALGSYMSKFTTETSQNFSTYQKGTGALSA-TGCVSOSNVYFPGVLOHESRD 949  
 Qy 981 MKKNS---YLMWTGLELNGCYTDTGTOIKISRDTGHRVTOQIKOSSKTTLYNDLNHIG 1037  
 Db 950 NKLISSGDDYRYSMSGLQRHKDKSFAHDHVSYDAEGRLVKTESSQVAFETDNVGRLT 1009  
 Qy 1038 SOWTDLATCHMLMTTVERDGLNRBGRKUQCDSSHTLDQQSMLKTQOLANTRVKGNT 1097  
 Db 1010 TTKDRTTLSQSLATKIEVDVDFBRIKRSLSDFSIQV-ITLSYTKNNOISORITSDCV 1068  
 Qy 1098 QRTEQYSDTSRRNLQNYKDCGAECPTDKGHSTVQNTYDIDYENITACHTFADGEHD 1157  
 Db 1069 MKNERVYDTSNORLUSOYDCEGEGSPVDTGGRVISQOITHYDQNGENIKLDNTYRDGET- 1127  
 Qy 1158 ATPKFANPTDPCQLTEVHHTHPDMDPNTRLKYDAGRVNINT-DNHGNTENTFTDGL 1216  
 Db 1128 VDHFHSQ-ADPTQOLRI---TSDKQJQELSYDANG--NLTRDEKGOT--LIVDQNR 1177  
 Qy 1217 --ONGOCSSV--YGYDPLNRLYSO-KDPTLDCBLYRTETMLUNEVENGEMIRL--- 1264  
 Db 1178 VQVKDSKGNLVCOYDADNLKTAQVLANGTVAHQ-YASGNVANVQLGDETITWLSSDK 1236  
 Qy 1265 -----RTGETIAQQRASKVLLTGSQSVLTSQDQNIQSEAYSAYSGKSTAND 1316  
 Db 1237 ORLHOSTKGKESVYQY-----GTHNSTVIASENTELMASTPYGFRSLI--- 1285  
 Qy 1317 ASIYGNGERADPVGVTGLNGYRSYDPTLMRHTPSLSPRGAGGINPNSYCLGDRN 1376  
 Db 1286 SLPGLNGAQVDPVTGWTFLGNGVNPYTCQGDPIN 1345

RESULT 3  
 Q44UTS\_PSESY PSESY PRELIMINARY; PRT; 1578 AA.  
 ID Q44UTS\_PSESY  
 AC Q44UTS\_PSESY  
 DT 13-SEP-2005 (TREMBLrel. 31, Created)  
 DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)  
 DE YD repeat.  
 GN Pseudomonas syringae 2044;  
 OS Pseudomonas syringae Pv. syringae B728a.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=205918;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RG DOG Joint Genome Institute;  
 RA Chain P., Larimer F., DiBartolo G., Copeland A., Lykidis A., Trong S.,  
 RA Nolan M., Gotsman E., Thiel J., Malfatti S., Lapidus A., Detter J.C.,  
 RA Land M., Richardson P.M., Kyprides N.C., Ivanova N.;  
 RT "Comparison of two complete genome sequences of Pseudomonas syringae  
 Pv. syringae B728a and Pv. tomato D3100";  
 PROC. NATL. ACAD. SCI. U.S.A. 0:0-0(2005).  
 [2]  
 RN  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Loper J.;  
 RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64;  
 DR STRAIN=B728a;  
 RA Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Feil H., Feil W.S., Lindow S.B.;  
 RL Submitted (MAY-2005) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL; CP000075; AAY37087.1; -, Genomic\_DNA.  
 DR InterPro; IPR006530; XD.  
 DR Pfam; PF0553; RBS repeat; 6.  
 DR TIGRFAMS; TIGR01643; RBS repeat\_2x; 2.  
 SQ SEQUENCE 1578 AA; 177054 MW; 514B576ABC10C9 CRC64

QY 252 PIGKNGILGQWITISMTAPEGGLKETVNYSNNNQGHFFPOSANLPUVPTUMLKVPGAGQP 311  
Db 234 LIREH---LCMISVETPGGRERIFYDG--GHOPPARARPLPRTRRTPGLQA 287  
QY 312 AIOEKEYSY---TSHNYVGGSNGIW-NNKNDLYGIMTE- YNYSSTESTRKYDKRGH 363  
Db 288 QMDVFSYQDSAGRBNFLFGAGDIAWEDNLQDMLYRIVGAELKEYSSTETLKVQAGKP 347  
QY 364 DQIWRRIERTVNMLTSCKQONGYIQTETAYTAALIGHNFSQPSQFOLRKTKTEF 423  
Db 348 DAERSIERTFNQFHLLTRETRTTRKLTVMQYDTRYTIVPQFQPPAPAYCOLPKEVQTR 407  
QY 424 -SADNSY-RSEIETTFDASGNPLTKVIKDKTOKIISPSTHWYYPAGEVDNCPPPY 481  
Db 408 LSSGSVPRISEMVSSTYDHGHLNLT---QANGMVEVS-- EWPAGE-DGCPPE 459  
QY 482 GFTTRVKKTIQTPDSEKFODPEKPIQYNSLGSOSHVTKIBERHYSATQIANS--- 537  
Db 460 GFVFRWMKSSKIPASDYGNAVPLVTRKIALPAVDGGCQLQOQWLAPOSETLLMQAPEB 519  
QY 538 ---TLEQINTDSEL---GRLIKQKETCTKGENGKT---YSVWHKFYTQDDTQS 585  
Db 520 YEQRLUTYEFDEPNALLGR-LKLQRYVIMGENATSTEVAYASTPDSYTGK- TVLQTV 576  
QY 586 HSITHDNFTIHIRSOVRSRTYGRLFSDTTKDTIVTQMSDKUGRLTRLNSGTP-YANT 644  
Db 577 QTIVGFGDAKKVLTLEDLSLANGEPELNRDNNDNNVNEIRYVUDSLRVVSETWAPNDIYAT 636  
QY 645 LTVDYELNNLQDDNRPFPVTTDYNQNLRFEDGAGRHVSQCLKDSDG---KFVYI 701  
Db 637 RSYEVQLCAKATDQTO---TLIDMKVKTCSHFDGLNRVTHERRDADPKARVDPQT 693  
QY 702 HTOQDXEOSHHTSTSVDYLNGRQTDPKVHSMSKSYDNGQIANTHWSYGVSEKIT 761  
Db 694 YAAYVDAGWGLMVNETEPDWL--GOODALKSVY---BYDWDGEQRCVTGADNIKTFEE 746  
QY 762 VDPI-TLATKQLOQNSNNNVQGKEVTTIPS---QPIQITLFDAGHQLQSCHTLR 815  
Db 747 TDPICTAQSGQGPVORSWIGTDGKCKSFETTWMLNQEPOTSERQDKAGTYSLSQHY 806  
QY 816 DGWDRVRKETDAI---GQTYIQDNYNVRQITLPDGTIVNRKXKAPFSTDLTIDIV- 871  
Db 807 DGLGRVVEEVGVNGGRRTTYYDFDRTESTUPSEAVRRSYASHSSEPDLSIGE 866  
QY 872 -NG- -ISLGQOQTFDGLSLRQTSQDGRGRVWATYSAGNDQCPSTVTPDCQFTHQYQPEL 928  
Db 867 HNGKSVIQLGQSFQDGLDLVSVTIGERKRKRTIVDVGLLQ- PRTVILPSKKEKIDYEYKPV 925  
QY 929 DD-----AVLQVASNEITQOFPSYMBTGALKAVA 958  
Db 926 GDEPWRVRQSBAAEPBDPTDTRVDPDKLQSGPGAAPARTAEGVAAADYTYDKQANLHSEE 985  
QY 959 EGQSLTPPIYPSGRJJKMENINDMKK---MSYLWTGLLENGYTDIGTOKISKRTDGRV 1015  
Db 986 GEIILDRREVSTTGJSKSEKRTSAGKTYDMVSYSQGULLSYTQVLQSGQVNRYDTCRIL 1045  
QY 1016 TQIKDSSKTTMLNYD---DILNRHTGSQVTDLATHGMLTVEFGNLNREIGRKUCDSSG 1071  
Db 1046 BETRLEGEVYSTPAYDVVFGRLLAKITTDQDSSRERTVAISLEYTDLGREVKRIF- DLDG 1104  
QY 1072 HTLDQDQSMWLTQQLANRIVYKINGVQORTQSYOSRNRRNQYKDGASCPTDKYGHSTV 1131  
Db 1105 VEQQMVQVYDQDVMWQRTSSEGTEILRDHNVDNGSRLTQVCGTRPFDVPGYTRL 1164  
QY 1132 TQNFTDLYGMITACTTPADGETDHATFKPANPDPQCOITEVHT-HPMDPDNRILKD 1150  
Db 1165 GQPFIFDGAQNLTEVTRP- DGGSNARY- FYEGGDPPVQSRVUTNHAYVPEINLRV 1222  
QY 1191 KAGRVNNTDNGNTENFTYDGLRQN---GOESVYGDPLNRLVLSQKTDCELY 1245  
Db 1223 PDGNL--DIDBAGRT- LKUDPLGLRLIEVGLTSAGIHQYQDQDOLTE- TRGGDRDRP 1277  
QY 1246 YRVMHKSGQEELVGLSGNRALPVRVYSPEGHGTILHYASCAYQMLSEVVD- DQG 178

---

RESULT 4  
Q883V9\_PSESM ID Q883V9\_PSESM PRELIMINARY; PRT; 1562 AA.  
AC DT 01-JUN-2003 (TREMBLrel. 24, Created)  
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE YD repeat protein.  
OS Ordericobusname=PSPT02239;  
OC Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonadas.  
NCBI\_TAXID=323;  
RN [1]  
RP STRAIN=DC3000;  
RC MEDLINE=22824015; PubMed=12928499; DOI=10.1073/pnas.1731982100;  
RA Buell C.R., Joardar V., Lindberg M., Selengut J., Paulsen I.T.,  
RA Gwin M.L., Dodson R.J., DeBoy R.T., Durkin S.A., Kolonay J.F.,  
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Hatt D.H.,  
RA Nelson W.C., Davidsen T.M., Zafer N., Zhou L., Liu J., Yuan Q.,  
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,  
RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,  
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhou S., Chatterjee A.K.,  
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,  
RA Bender C.L., White O., Frazer C.M., Collier A.;  
RT "The complete genome sequence of the *Arabidopsis* and tomato pathogen  
Pseudomonas syringae pv. tomato DC3000."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).  
DR EMBL; AR01653; AA055755.1; -; Genomic\_DNA.  
DR TIGR; PSPT02239; -.  
DR InterPro; IPR006530; YD.  
DR Pfam; PF00593; RHS repeat; 6.  
DR TIGR; TIGR01663; YD repeat\_2x; 3.  
KW Complete proteome.  
SQ SEQUENCE 1562 AA; 175713 MW; 8DC10DA1BFE37B1 CRC64;

Query Match 18.2%; Score 1614; DB 2; Length 1562;  
Best Local Similarity 30.3%; Pred. No. 1.2e-78;  
Matches 490; Conservative 239; Mismatches 700; Indels 190; Gaps 51;

QY 14 MSDNNEFFQANNFTISAVGGVDPRTGLNQITLGHVGNHNGPFTLTLTSYPLKNT 73  
Db 1 MTTTSVSHNAFNFMVSYLQSGVDPRTQTVSINLPEVKSGLGPVPLVSLNPLVQ 60  
QY 74 DIGRIGFNGLSYDRKNSLISSTGEMYKVI-BTDKTVKLQOKKLDRPEKDLKENC 132  
Db 61 DSGCIGLQWNLQLSdyPGeTRIVSGSGETVKGECISLGQWLMPEKKDHFVYQ-DTR 119  
QY 133 YRILHKGDIKEVLTGCFNNAFNLKPKCUNPAGHAIYDWNFEATQPRLNRYDLDGH 192  
Db 120 YRVMHKSGQEELVGLSGNRALPVRVYSPEGHGTILHYASCAYQMLSEVVD- DQG 178

QY	193 DIPLLNLBEGLIKITLTUFGQKREGYTERLPLNQLNSTHFLSLGENPLTWSGRYTP 252
Db	179 VILTTTRD---STSVRLLYGAPKADEFMILSSNRNARIELPTANKASWRFTSI 234
Db	253 IGRONCGWNTSMAPGGKEBTWVNSNNNOCHHFRSANTLPVPLVTLQKQVAGOPA 312
QY	235 IRGHC---IASVUTPVGKGHEDFVFIQD--GHOPPSAGREPLPRTRHLTTPGFLQPE 288
QY	313 IQAEVSY---TSHVYGGGSNGI-W-NNKLDNLYLGIM- -TPYNGSTESTRKYKREGHD 364
Db	289 VDVRYAYKGAGREHNPFLGAQDIAWEDNGLDNLNLYRVLGAAPYKSTETARVNDD--- 345
QY	365 QIVRERTYNNYHLITSECKOQONGYQTTETAYAATIGHNIDQSPOSQFQDLRKTETWR- 423
Db	346 -VRSERVFQPHLAETTRONLSLENDTRYVIEGKPPDLQPNYCQLPKERVTWL 404
QY	424 SADNY-RSBETTETDESPENPLTKVICKYOKLUSPSTIWEPYYPAGEVDNCBPERPYG 482
Db	405 SPDGSUPRTBVSSDSYGMILAQOANGVETS-----EWYSVSGE-DGCPPDPG 456
QY	483 PTRFKKILQTPTDPERKDPKFIQRY---SLIGSQSHVTKIE----BRHSATQ 533
Db	457 FVRTLAKASTVPAODQGDYGHALVLYTRYKALPAGLGSQNLWLAESETLQQTGEK 516
QY	534 LLNSTLFOYNTDKS---ELGRLLKOTECTKGENGK---YSHWHKPTTYKQDDTQOSH 586
Db	517 ELQQTYYTYEDNPDAQFQGRIRHOSVTLGSLSTTDYRDQDPD--DQTVQQT 573
QY	587 SITTMHDNFTHRSOR---SRYTGIFPSDPTKDVTQMSYDVKGRLLTRTN-SGPVYA 642
Db	574 QIVN---GFDMNQKVURLEHSLIFTGEBLNRRNDNDVETRYDNLRRVSETVSPNEYK 631
QY	643 NTLYDYBLNLQDDNRPFPVITTDYNGQNLNEFDGAGRHSQCLDSGDKFPTIH 702
Db	632 ATRHVEYOLGAVKTPDABQBLF--DVKYVOTTSRDGLGAVIYEARADADNPVRRRLD 688
QY	703 TQO---YNEQGRAHTSTSDYLNGRQQNDPDKYHLSMRSYDNGQIANTHWSYGV 757
Db	689 LRQTYEAYDAGDKVETSYDWL-----DQQKGALNTNPYBDWDQQLSVTGFDPGT 741
QY	758 EKITDPTIPITAT----KOLQSNMNWQTCKEVRY-TSQSQPQITLFBAGHQSC 811
Db	742 TIEQTDPGVQTOASNSPIORVTTESNGLTSBSEVTEWLNTDEPTMSVRDLWNSBEPV 801
QY	812 TLTR---DGADRVRKETDAI---GOCTIQYDNYNMRVIQITLPDTIVNKKVAPFSTD 865
Db	802 SLSRVQYDGIGRLVKEVNGSLBIRESTTYGDPVERVANTLPDGAVVRVYAPHSGRD 861
QY	866 ITDIRV--NGTS--LGQQTENGSLRILUTOSOPGGRWMTYASAGNDQCPSTVTPDQSFTH 921
Db	862 PAWIGDHNGKSSVIGEOKFGLDRIVTSITGGRERBLSYTSDLMO-PKTKLUPSGRQD 920
QY	922 YQOBELDDAVLO---VASNRIIQFSNPNTVPGALKAVAEGSOSLTPYPSGRBLK 975
Db	921 YDYGDRGDLKRTQSDTWRKLADYDTPONARLIGSSEQGBLORVYSTSLS 980
QY	976 ENINDM---KKMSVWTLRGLENGTDLTIQKSRSRTHGRVTOIKDSSIKTTANDDL 1032
Db	981 EQRISQGIENTMHWYNSRIGLPLSYDVLGEOBQSLTYDFFCRAQTSLGEVUSDFWTF 1040
QY	1033 NRHTGSQVTDLATGHMLTTVEFDGLNREIGRKLAQDSSGHTDIOQSWLKTQOLANRIV 1092
Db	1041 GRTASSIATDSSNSQCVVSLYDAGREAOITP-TINGANQOMQVQYDDQWMKRTLS 1099
QY	1153 GTRDMMATPKFANPTPCQLTTEVHHT-----PDMPNIRLKDKAGRVIN 1197
Db	1159 GGRNHPARY-FYEGIPAQLTRVNTQALAWNWRALIPIVNUKNDSTYPPETRUYTDPGML-- 1215
1198 ITDNNGENTENPTYDTIGRL---QNGCSV-YGYDPLNRLYQSQTDTDCELYRTERMVN 1253	
RESULT 5	
ID	QTNWU2_ CHRV0 PRELIMINARY;
AC	QTNWU2;
DT	01-MAR-2004 (Tremblrel. 26, Created)
DT	01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)
DB	Nematicidal protein.
GN	Name=rnba; Order=locusNames=CV1887;
OS	Chromobacterium violaceum
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Chromobacterium.
OX	TAXID=336; [1]
RN	NUCLEOTIDE SEQUENCE.
RP	STRAIN=ATCC 12472 / DSM 30191;
RC	MEDLINE=2282880; Pubmed=1450782; DOI=10.1073/pnas.1832124100;
RA	Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T., Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R., Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.R., Astolfi-Filho S., Azevedo V., Baptista A.J., Batatas L.A.M., Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S., Bordignon J., Brigidio M.M., Brito C.A., Brocchi M., Burity H.A., Camargo A.A., Cardoso D.P., Carneiro N.P., Carraro D.M., Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O., Crecynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L., Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A., Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R., Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B., Grattapaglia D., Grisard E.C., Hanna B.S., Jardin S.N., Laurino J., Leoi J.C.T., Lima F.H.A., Loureiro M.F., Lyra M.C.P., Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S., di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M., Nascimento M.Z., de Oliveira M.F., Oliveira J.G., Oliveira S.C., Paixao R.F.C., Parente J.A., Pedrosa P.O., Pena S.D.J., Pereira J.O., Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.H.R., Potrich D.P., Ramalho-Neto C.B., Reis A.M.M., Rigo L.U., Rondinelli R., Santos E.B.P., Santos F.R., Schneider M.P.C., Seuanez H.N., Silveira A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C., Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L., Souza R.C., Steffens M.B.R., Steinidel M., Teixeira S.R., Umenyti T., Vettore A., Wassem R., Zaha A., Simpson J.J.G., "The complete genome sequence of Chromobacterium violaceum reveals remarkable and exploitable bacterial adaptability"; Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR	PFAM; PF0553; RHS repeat; 6; TIGRFAMs; TIGR01643; YD repeat 2x; 3.

KW	Complete proteome.	SEQUENCE	1385 AA;	152878 MW;	BAAB59C291CBBE CRC64;
Query Match	18.0%; Score 1598; DB 2; Length 1385;				
Best Local Similarity	32.2%; Pred. No. 7, 3e-77;				
Matches	460; Conservative 222; Mismatches 575; Indels 170; Gaps 52;				
QY	14 NSDNNNEFTQAMFTSAGGGUPRGLNQITLGHIVNGNLGPIPLTYSPE-N 71				
Db	1 MADNTGIGSQAMNPSAVANGYDPRGFLSPNFDLGLSANSOGLTALPILRAYPLAAG 60				
QY	72 KTDIGIGIGPNGLSV-YDRKNSLISLSTGENYKVIEHDKT-VKLQOKKLDRFLPFDLK 129				
Db	61 ESSDGIGKGSNGFPIKDRFLRLLIASGERYRLIDEDSAFVPIKQLKDVRVHRTVA 120				
QY	187 DLDGDHPLMLNBYQSL-IKTLTRP-GKEGYRTLEPLRNPQMSIHFNSLGNP 243				
Db	179 DE--EKGTLKIDPVNADGERVATIQWPSDEKVALEYFQNGYLHVNKSLSONGD 235				
QY	244 LWTWSFGTPTIGK--NGLIGOMTSMTPRGGLKETVNTSYNNNGHFFOSANLP-----V 295				
Db	236 VEWLTGIFTDSKVADEVGILLNETAPTGLAQVRVB-----PLCMKQGERSDFG 287				
QY	296 LYTUTLMQVPGAGQPAIQAEVSYTSHNYG---GGSGNGTWKKNDLNGLMTEINYG 350				
Db	288 LPAVVLISLVPBAGQAPAINTHYEYSPANTLYGASFKQSOG----GADELFDIQVPTYQ 343				
QY	351 SRSRSRYKQEGHDQIVRIERTNNYHILTSKQKQVQIQTETAYAIIHSNFQSOP 410				
Db	344 SIEKLUKLSLNPKPKRPTVTKNNPHILVSEREGACVPROBTAVPAVKQSYEAQPA 403				
QY	411 QPOLPKKTTKETRSADNSYRSEETETPDESENPLTKVTKDKCTQKLISPSTHWEYPPA 470				
Db	404 TFLQPVVRQTMWSMEEAGRSRRESIS-FEYDDAGQLIKOMSD-----GSITVLEYPPA 455				
QY	471 GAVDNCPCEPYGPTRFK-KIOTPYSEPKDPEKFQIQRSLI---GQSHWTIKI 524				
Db	456 GBRGHCPCADAEGRGRYLUKSKTVPEVPSVSABYDEMAMREHYEVFRSRTRPGSAMHAETLQQ 515				
QY	525 BERYSA-----TQLMSTLROY--NTDKSERGRLIKQTECTRG 561				
Db	516 TVSHYAGMPGPKARAMILGGKPSWEAQPROLAKESTYDAPAQKHGRIKRKIAVVG 575				
QY	562 ENGKTYSVVHKFETY-----KODDTLOSHSIATHDN-----FTHRSQVRSSRTG 607				
Db	576 EDRTPYENVQDFVFEPVRSGNREVALKQFVSVTKEDEPLEKKDKGOKRSVSTRVLSVTG 635				
QY	608 RLFSDTDTKDIVTQMSTDKLGRLLRTLNSG-TPYANITYDYELNNIQLDDNRPPFVIT 666				
Db	636 RLUSETDVLGNTVAYGPOLGRKTOAHPDLYKAYRAJARPARWDYLWPSPK--NGTPAMAH 693				
QY	667 TDVNGNORNERFAGRAVSOQLKSDSDGKFTHTQOYDSCGRHTSTSYDLTNGRQ 726				
Db	694 TDALGNQRTSHDGLGRMIREEACDRDGGLGKVTWRTHYDEAGRQARTITVTVHDL-R 751				
QY	727 QTPDPKVHLSMSKSYDWGQIA-NTMWSYGVSXKTVPIPLT-ATKLOQSNSNNVQGK 784				
Db	752 EGKPVTLKTEREDWSDGOLSVERELETBLGLASRQEIDPIAOFVATWAGTR--CSAK 808				
QY	832 TIVQDYNTRNVIQITLPGTIVKRYKAFASTDPLTILTDIV--NGIS--LGQTFDGSRL 887				
Db	809 YMFYFSKSHDLRRRIVJ--ATHLESWSWDAFDPKYVASWNGWDGAHRRAATENSHA 865				
QY	888 TQSQDGGRMWAVYIAGNDQCPSITVTDQFQHYQPELDDAVLQAS-----N 938				
Db	926 KRTESGRPTREYASDASSRTVCGDGRQVAYTDRLGEALKSYAAKDPHQLGVS 985				
QY	939 BTQOFSNPVPGCALKK--VAGQSLSTPIVYPSGRKMENTNDM---KNSYVLWLR 991				
Db	986 PIQOTSYLLPMLLHEBEVSGAQSAWD-RWPSPGRLR-EETHDIRSGKKQKOHYRSLT 1043				
QY	992 GLENGVDLDTGTQKISRDT--H-GRVTOQDSKTIKTLNDLNRHHSQYDTLGH 1047				
Db	1044 GLUEGADDIGDAHARSYETAAHVGKLTIDAATVTLAYDGLORUCSWTARD-GRGH 1102				
QY	1048 MLTTIVEPDGLNKBIGRKLCDSGHTLIOQSWLKTQOLANLIVKNGVLTQEQYSDS 1107				
Db	1103 ALATTLEFDSDAERETKTLAASAAEATLSQLQWVYNGQOLHQKRSBGKPFDETVDYD 1162				
QY	1108 RARLNOYKTCGDAECPDKYGHISVTONFTYDLYGNTIACHTFADGTEDHATPKFANPTD 1167				
Db	1163 RURKVOYAASGPGLPKDQAYGNAIRQKRFEFDAFNIRKCTTLDGGSBNVSYLFENPAD 1222				
QY	1168 PQLTEVHHTHD--MDNTRKXDKGVRVINITDNGNTENFTYDGL--QNGQGS 1222				
Db	1223 PQLTKVNTNSALDKGYPAEILKQYDQGLR--DEAG--RRLSYDALGRARVEGGGS 1278				
QY	1223 V-YGDPLNRLVLSQKDT--LDCELYTRETMVNE--VRNG-----EMIRLRTGETI 1270				
Db	1279 ASYGDYADHRLICQRCVETSGMDHRLYTRANRJNEWMTSGGPGADDVRLVYAGSC 1338				
QY	1271 INQOR---ASKVLTGTDSDOOSVILTSKDKNQSBQAYSAYGRKST 1313				
Db	1339 AJOVNEGDDGSVAALMGTDGKGSIVSOREGGAKHYATPYG-HOSS 1384				
RESULT 6					
QY	OTN4A7_PHOIL ID QTN4A7_PHOIL PRELIMINARY; PRT; 1775 AA.				
AC	QTN4A7;				
DT	01-MAR-2004 (TREMBL) 26, Last sequence update)				
DT	01-MAR-2004 (TREMBL) 26, Last annotation update)				
DB	Complete genome; segment 9/17.				
GN	OrderadociaNames=pJL242;				
OS	Photornabidus lumenesens (subsp. laumondii)				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.				
OX	NBBI_TAXID=141679;				
RN	[1]				
RP	NUCLEOTIDB SEQUENCE.				
RC	STRAIN=T701;				
RM	MEDLINE=2257627; PubMed=14528314; DOI=10.1038/nbt0886;				
RA	Duchaud B., Rusniok C., Frangeul L., Buchrieser C., Givaudan A., Raoult S., Bocs S., Bouraux-Blude C., Chandler M., Charles J.-F., Dassa B., Derose R., Deruelle S., Freyssinet G., Gudrault S., RA Medigue C., Lanois A., Powell K., Siguer P., Vincent R., Wingate V., RA Zouine M., Glaser P., Boenare N., Danchin A., Kunst F.; RT "The genome sequence of the entomopathogenic bacterium Photorhabdus luminescens.", NAT. Biotechnol. 21:1307-1313 (2003).				
RL	DR EMBL: BX571867; CAE1816.1; -, Genomic_DNA.				
DR	DR Photolist; Plu2442; -.				
DR	DR InterPro; IPR005530; YD.				
DR	DR Pro3; PR0593; RHS repeat; 6.				
DR	DR TIGRFAMS; TIGR01643; YD_repeat; 2x; 3.				
KW	Complete proteome.				
SQ	SEQUENCE 1775 AA; 200136 MW; 0621A9FC06DB8801D CRC64;				
Query Match 17.8%; Score 1583; DB 2; Length 1775;					
Best Local Similarity 28.2%; Pred. No. 7, 3e-77;					
Matches 51; Conservative 280; Mismatches 727; Indels 294; Gaps 65;					
QY	22 TQANNTSAVSQGDPRGTLQITLGHIVNGNLGPIPLTYSPLNKTDIGFCGF 81				
Db	19 SQATNFTGAMM3GVDPRGFLPFIQPTQIOLQKPLTQVYGLP 778				
QY	82 NEGLSVDRKNSLISLSTGENYKVIEDTKV-----KLQKQLDNLRFPEKDJKENCRI 135				

Db 79 SDNFTRYPTOTVQLTATGEIYHRVEKANEVPGQAWTPHAKPAHKVKR--KDAFWV 136  
 Qy 136 IHIKSGDIEVLGTFNNNAFDLKVPKPLNAPGHAIYIDNNEATOP---RLNRYDDLG 191  
 Db 137 LYKGSREKLTQ--DRAMPVAVUVEIYARSGHCKLVKNSPFLHMYNTWQLMHEVCDADE- 194  
 Qy 192 HDIPLLNLBYQGLIKTILTLPSOKEGETRTEFLNROLNSTINPSLGENENPLTWSCGT 251  
 Db 195 --TLLKADLATTEKIEFTVWPSPESTVTLMNTDLQTVISAS---ELJWHLEYE 246  
 Qy 252 PICKNGILGQWITSMTAGSGLKEFTWYNSNNQHHP----- 288  
 Db 247 TEGAHKNI---LTKVTTSGLIEKVVY--HETGHTLPTPKC1QYPNAWGPSTRODKS 301  
 Qy 289 -QSANLPLPYVTLMKOPGAGOPAIQARYSYTSHNYVGGGSGIWN---NKUDNLGL 343  
 Db 302 STRATVENVHPYVTOHHLTAGSGSDQVTRVYFSPENFGOKONMDIPLQODNAYA 361  
 Qy 344 MTEHYNGSTESSRKYDKDCHQDVIERTYNTNHLTSCK---QONGYIOTTERAY 397  
 Db 362 NSEKYKISTEVREYNEKR---YCIRHYNKFHLLVBETTEVTPSRPQKGETIKY 416  
 Qy 398 YA1GHNED--SOPSOFOLPKTKETMSADNS---YRSETETTEDBSMLTKIK 450  
 417 YADVGSKSPDNKQGPQPLMPNTVETIWINPESASSATORRKETTOMENAOQNLMISWT 476  
 Db 451 DKKTQKISLSPSTHWYVPPAGEV---DNPPEPYGPTFRVKI-IOTPVD-SFKDDEK 505  
 Qy 477 DNTTKT-----TYARDGEETDTGCPAERPGPERIFKEAVAEAPSPLITKITLR 529  
 Db 506 PIORYSLIGOS-----HVTLKIERMYS---ATOL---LNSTLFOINTD 545  
 Db 530 KVINYKSYDTSPPKNNQASVVKSMVPLISLETHYSRDCDCHLEKVKINTVSFYENTQ 589  
 Qy 546 KSEGLRLIKOTECTKGENG-KTYSVWHRFTYKODDITQOSHSTITHONF---THRS 599  
 Db 590 NAPFLNRYVEORNSYLPGENGNTQSYTEDYSWSNENKNGASCCKTTSGKCCGTPPVSH 649  
 Qy 600 QVRSYTGRLFSDTDKTDKIVTOMSYDKUGRALLTRTLNSCTPPYANTLTWDYELNLQDDN 659  
 Db 650 QYNSRSTERBLIPOKDAQDINTFQYDPTGLHLLISTINADAVERKTYKAWSTSNKR--- 705  
 Qy 660 PPEVITTDVNGMOLNERDAGGRHVSDOKLSDPDGKREYTMHQYDPROGHRTSTYD 719  
 Db 706 -VVVTQDIHENRYITEMDGLGRPLKSYSPACYQGKQDFMERYQINPLGLCQALSCD 763  
 Qy 720 YLTNGRQDTPDKHLSMSKSYDNGQANTHUSYGVERKTVDPITATKOLQSNN 779  
 Db 764 HML-PENKSEKCKCSSMTVSLBDDWGHNSRRLYRSDGTSVNTFDPKMTSEHORVSNDES 822  
 Qy 780 VOTKEVNTYTPSQQPIQTLFDEAGLHOSCHTLTROGDVRKETDAIGQCTIQDNY 839  
 Db 823 OSSGILRTYNTQEQQITATEPLTTSRIGCWHLRDELGR-L-VSINANGNTLLDAF 881  
 Qy 840 NRVQTQITLPGDTVNRYKA-PFSTDLLTDIRN-GISLQOQFDGSLSRLTOSDGRV 896  
 Db 882 DRVJKQTPADGITISMAYENGSVRMSATPLGINQTPVILGTQILDGLGRVIDMESGK 941  
 Qy 897 WATVYSGASNDQCPSTWVITDQO-----PHYQYQPELDLQLOVAS--NETTQFSYN 947  
 Db 942 IKUDYE-GASPVPDVTVVKGPSTGGDKDVTUHVYEPKLNAAVTKITAGQDVFDFKD 1000  
 Qy 948 PVTGALLKAVE-----GOSLTPYPPSSRLME---NINDMKKOMSVLTLR-G 992  
 Db 1001 PKYGLLITEREYKNDMESISKLIPDTSGLASEKLVSKTNKRNAYMHKTCVS 1060  
 Qy 993 LENGYDTLGTG-----I0KISDTHGRVQIKOSSIKTLYNDLNRHGSQVTDL-- 1043  
 Db 1061 FAMRPFTAICSTINGPRVINKSYDNGQLETVRNDIEEINLKDQKLRHVCOSTEVFD 1120  
 Qy 1044 ATGMLTTTVEFDGUNREIGRK-----LCDSCHT-LDQSWLTKOOTOLANIVKNGV 1096

Db 1121 TINNKITSTLTDGFGRKTRTIRRHLGKNNKGITDHISQTFNEQDKITKEL-LHGT 1179  
 Qy 1037 LQ-RTEQSYSDSNRNLQYKCDGAECPTDKYHSIVQ-NFTYDVGNTIACHTTDG 1153  
 Db 1180 KOJSKEVVTYKHGGLETYL----MELVBEREITQCSDYKDHGNITO-HSITTEG 1232  
 Qy 1154 TEDHATFKEFANPFT-DPQOLTEVPH-TFPDMPNIRUKDAGRVINITD-NHGNTENFT 1209  
 Db 1233 KITSTTYWNGNMQDPCQOLIDVSTCIGNSSSLRFTYNGORGALVGCENDENTKIRWT 1292  
 Qy 1210 YDTGURGJLNGQSV-----YGDPLNLV--SOKIT-LDCELYFETMLVN----- 1253  
 Db 1293 YDSIGRIDDTVDALKVETRYLFDATWKLKIKRSBRKGTPYHDLSTCSNLSLHDNTYFG 1352  
 Qy 1254 EVNGEMR-----LARTGCTIAQ--QRASKVLTGDSQSOSVILTSKQNSQ 1302  
 Db 1333 EKRINAADRKYKNGGICLIGFSOPCHOPTSASVRYETATDGKSVIATFOGEDQH 1412  
 Qy 1303 AYSAYGKHKSTANDASIG-----YNGERADPVGVTIANGYRSYDPTIM 1348  
 Db 1413 AVSPWG---VTRQAMVPAQGQPHNTAEEPFENGQEWDTASASYVINGTAYRPLM 1469  
 Qy 1349 RFTPDSLSPFGAGGINVSYCYLGDPINRSDEGHLSSQWATGIGMGIAGLITATGGM 1408  
 Db 1470 RFTAPDSHSPFGAGGINAYCQGDPVNLNDPSGHISCGWMANITGGIGLLAPFTGG 1529  
 Qy 1409 ALAAGGAAALASTSTALAFGLSVSPDTSIVSGLADEASPKAASSILGMVSMGCAA 1468  
 Db 1530 SLELGLGVNARGLT-----ALDAASGVTAISAGLAENKNPETSRLGKMSLGLG-- 1579  
 Qy 1469 GLAESAIGKGTKA-----THLGR-----AEDGENALLKSTB--- 1502  
 Db 1580 -LPSMVGGYSLAQWNVWRLTSFRTRHYPHSILGEUNLSRSKSDWNRKSLSGNW 1637  
 Qy 1503 -----SSRIKMG---VTRSLDRBIVRNBEGOKDHERGYTUNFMGKGEBAI-----L 1547  
 Db 1638 HSEVGLNGNRTIWGSDTKRGLD--IKYPLEQISRRENG--DIVLSSGSHVQNGDWL 1692  
 Qy 1548 VHDKDQGLYH-----TEGNKHMKKGPYTRHPTPQLVDLKDNNIVDLYQGGDLPV 1598  
 Db 1693 INGSRRGSLIHFPPFKSMTVYGGSKWRTVHVN----LATMSEIDFGTLLANGNHSI 1747  
 Qy 1599 HLLSCYGGSSGA 1610  
 Db 1748 ILGVCYGRNDQA 1759

RESULT 7  
 QSS6J1 COXBJ  
 ID QSS6J1 COXBU PRELIMINARY; PRT; 773 AA.  
 AC QSS6J1:  
 DT 01-MAY-2000 (TREMBrel. 13, Created)  
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Coxiella burnetii.  
 OG Plasmid OppD.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;  
 OC Coxiellaceae; Coxiella.  
 OC NCBI\_TaxID=777;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.

STRAINER140;  
 RA Radomski K.U., Willemse H., Lautenschlaeger S., Jaeger C., Baljer G.;  
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.  
 DR InterPro, IP006530; YD\_Plan; PR0559; RBS repeat; 6.  
 DR TIGRFAMs; TIGR0143; YD repeat\_2x; 2.  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 773 AA; 83759 MW; 4C5FC0481CCAC4 CRC64;

Query Match 17.5%; Score 1553; DB 2; Length 773;

**Best Local Similarity 44.6%; Pred. No. 8.6e-76; Matches 342; Conservative 119; Mismatches 275; Indels 30; Gaps 17;**

**QY 741 YDNWQIANTHWSVQSEKITNDITLAKLOSNNSNNVQTKEVVTPSQQPIQTL 800**

**4 YDSNGQNLHTVPSQYQEVYDPITRAT--LQPESSOKQOLTEENLAGLPIKTYQ 61**

**Db 801 FDEAGLHSCHTLTRDGDRVRKETDAIGCCTIVQDNTNRVQITLPDTINRKYAPF 860**

**62 YDSGQTEQSAYHSDGLQQLRKTDELQGQVTLVEYDHFGRVQTLLPENTIQSYAH 121**

**QY 861 STDTLITDTRVNGISLGQCPFGSLRITSQDGRVRWAVTYSACNDQCPSTVTPDGSFI 920**

**122 STASLITGQSVNFSMGNQTFDSLRLTETTSGGRTAPSYYENAS-SVPAVTAFTGTV 180**

**QY 921 HQYQPELDALVQVASNITIQOSYNSYNTPGALKAV-AEGSQLTPYPSRLKOME NI 978**

**181 SYETKLGNVAKKISAPELKTDWYDALGAMTSATOAGMIRQMTYPSGLKNETM 240**

**QY 979 ND---MKKOMSYLTURGLLENGTDLGTTQKISRDTGRVTKQDSIKTTNDDLRH 1035**

**241 PGAAQKSTAYTSLAGAPOSYTDFGVVQRYDHEGRGIRDIFDVLGDAFGRP 300**

**QY 1035 IGSQVTDLITGHMLTTVEFDGLREIGRKLCDSSGHTLDIQSOWLKTQOLANRIVLN 977**

**301 TKQOATDKKTGAVLSTLTYDLDLREBKRISASQSVLVIQEIQYRNHLKERITQR 360**

**Qy 1096 VLRQEQQSYTSRNLNQYKOGAECPTKQGHSLVQTMFTDLYGNITPACHTFADGE 1155**

**361 TTURKEVFAVDSRNLRLIBTYCNGEARPODPYKAIHQRTSYDAGNMKTQDFSGG-R 419**

**Db 1156 DHAFKFKANPTDPCOLTEVHHTPDMNDNIRLKDKAGRVINUTDHGNTENFTDILGR 1215**

**420 NTATIVY-SAIDPTQLLKVNNDHSYDPIKEITLTDKAGMR--DEAGRT--LRYDAIGR 474**

**QY 1216 LQ--NG---QGSVYGYDPLARLVSQKT-DTLCBLYRETMVNEVR--NGEMIRLRTG 1267**

**475 LQQVNGAGAKGQYADALNTLVSQVQDPIYDLYRADDLVSGARRDGSQTRIVKN 534**

**Qy 1268 ETIAQ--ORASKY-LLTGTDSPSOSVILTSKDNLSQRA-YSAVGKHKSTANDASIGY 1322**

**535 GGCVQGCTKQGSNTSRITTTNQCSVLVSSEGNAHPDQCYTRGRTQPTEPVSLGR 594**

**QY 1323 NGERADPVSGVTHLGNQSYRDPMLRFPTPDSLSPPGAGGINPSYCLGDPINRSDEG 1382**

**595 NGERLDPVSGTYHLLGNYRANPILMRFNCDPSWSPFGAGGINPYACQGDPPINRVDNG 654**

**QY 1383 HLSWQAWTGTGIGMIGAGLILTTATGGMMAIAAGGIAMAIASTTALAFGALSVSDITSI 1442**

**655 HLSWQAEGLGIGLVGVULAVFTAGTSIANGAISAAEASAISLUVGTLGVAADVSI 714**

**QY 1443 VSGALEDASPKASITLGWVNGMGAGLAR---SAIKGJPLATHL 1485**

**715 ASGALEDANPOASATLGWISLGIGGGPAGVSGLATAARAGKKLISGL 760**

**RESULT 8**

**O5880 COXBU ID 052880-COXBU PRELIMINARY; PRT; 774 AA.**

**AC 052880; DT 01-JUN-1998 (TREMBrel. 06, Created)**

**DT 01-JUN-1998 (TREMBrel. 06, Last sequence update)**

**DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)**

**DB Hypothetical protein orf\_774.**

**GN Nameorf\_774; OS Coxiella burnetii.**

**OG Plasmid QRS.**

**Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;**

**OC Coxiellaceae; Coxiella.**

**NCBI\_TAXID=777;**

**RN [1] (TREMBrel. 31, Last sequence update)**

**RP NUCLEOTIDE SEQUENCE.**

**RA Lautenschlaeger S., Jaeger C., Willemse H., Baljer G.;**

**RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.**

**DR EMBL; X15898; CA075841.1; -, Genomic\_DNA.**

**DR InterPro; IPR006330; YD.**

**DR Pfam; PP05533; RBS repeat; 6.**

**DR TIGRFAMs; TIGR0143; YD repeat\_2x; 2.**

**KW Hypothetical protein; Plasmid.**

**SQ SEQUENCE 774 AA; 83774 MW; 918A533D7991BDB CRC64;**

**QY 741 YDNWQIANTHWSVQSEKITNDITLAKLOSNNSNNVQTKEVVTPSQQPIQTL 800**

**4 YDSNGQNLHTVPSQYQEVYDPITRAT--LQPESSOKQOLTEENLAGLPIKTYQ 61**

**Db 801 FDEAGLHSCHTLTRDGDRVRKETDAIGCCTIVQDNTNRVQITLPDTINRKYAPF 860**

**62 YDSGQTEQSAYHSDGLQQLRKTDELQGQVTLVEYDHFGRVQTLLPENTIQSYAH 121**

**QY 861 STDTLITDTRVNGISLGQCPFGSLRITSQDGRVRWAVTYSACNDQCPSTVTPDGSFI 920**

**122 STASLITGQSVNFSMGNQTFDSLRLTETTSGGRTAPSYYENAS-SVPAVTAFTGTV 180**

**QY 921 HQYQPELDALVQVASNITIQOSYNSYNTPGALKAV-AEGSQLTPYPSRLKOME NI 978**

**181 SYETKLGNVAKKISAPELKTDWYDALGAMTSATOAGMIRQMTYPSGLKNETM 240**

**Db 979 ND---MKKOMSYLTURGLLENGTDLGTTQKISRDTGRVTKQDSIKTTNDDLRH 1035**

**241 PGAAQKSTAYTSLAGAPOSYTDFGVVQRYDHEGRGIRDIFDVLGDAFGRP 300**

**Db 1035 IGSQVTDLITGHMLTTVEFDGLREIGRKLCDSSGHTLDIQSOWLKTQOLANRIVLN 977**

**301 TKQOATDKKTGAVLSTLTYDLDLREBKRISASQSVLVIQEIQYRNHLKERITQR 360**

**Qy 1096 VLRQEQQSYTSRNLNQYKOGAECPTKQGHSLVQTMFTDLYGNITPACHTFADGE 1155**

**361 TTURKEVFAVDSRNLRLIBTYCNGEARPODPYKAIHQRTSYDAGNMKTQDFSGG-R 419**

**Db 1156 DHAFKFKANPTDPCOLTEVHHTPDMNDNIRLKDKAGRVINUTDHGNTENFTDILGR 1215**

**420 NTATIVY-SAIDPTQLLKVNNDHSYDPIKEITLTDKAGMR--DEAGRT--LRYDAIGR 474**

**QY 1216 LQ--NG---QGSVYGYDPLARLVSQKT-DTLCBLYRETMVNEVR--NGEMIRLRTG 1267**

**475 LQQVNGAGAKGQYADALNTLVSQVQDPIYDLYRADDLVSGARRDGSQTRIVKN 534**

**Db 1268 ETIAQ--ORASKY-LLTGTDSPSOSVILTSKDNLSQRA-YSAVGKHKSTANDASIGY 1322**

**535 GGCVQGCTKQGSNTSRITTTNQCSVLVSSEGNAHPDQCYTRGRTQPTEPVSLGR 594**

**QY 1323 NGERADPVSGVTHLGNQSYRDPMLRFPTPDSLSPPGAGGINPSYCLGDPINRSDEG 1382**

**595 NGERLDPVSGTYHLLGNYRANPILMRFNCDPSWSPFGAGGINPYACQGDPPINRVDNG 654**

**Db 1383 HLSWQAWTGTGIGMIGAGLILTTATGGMMAIAAGGIAMAIASTTALAFGALSVSDITSI 1442**

**655 HLSWQAEGLGIGLVGVULAVFTAGTSIANGAISAAEASAISLUVGTLGVAADVSI 714**

**QY 1443 VSGALEDASPKASITLGWVNGMGAGLAR---SAIKGJPLATHL 1485**

**715 ASGALEDANPOASATLGWISLGIGGGPAGVSGLATAARAGKKLISGL 761**

**RESULT 9**

**O4ZUT9\_PSESY ID O4ZUT9\_PSESY PRELIMINARY; PRT; 1682 AA.**

**AC 04ZUT9; DT 13-SEP-2005 (TREMBrel. 31, Created)**

**DT 13-SEP-2005 (TREMBrel. 31, Last sequence update)**

**DT 13-SEP-2005 (TREMBrel. 31, Last annotation update)**

**DB YD repeat.**

GN ORPNames=PBYR\_2040;  
 OS Pseudomonas syringae pv. syringae B728a.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas;  
 OX NCBI\_TaxID=205918;  
 RN [1] NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RG DOB Joint Genome Institute;  
 RA Chain P., Larimer F., Di Bartolo G., Copeland A., Lykidis A., Trong S.,  
 RA Nolan M., Goitman E., Thiel J., Malatti S., Lapidus A., Detter J.C.,  
 RA Land M., Richardson P.M., Kyrides N.C., Ivanova N.;  
 RT "Comparison of two complete genome sequences of Pseudomonas syringae  
 pv. syringae B728a and pv. tomato DC3000.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 0, 0-0(2005).  
 RN [2] NUCLEOTIDE SEQUENCE.  
 RC STRAIN=B728a;  
 RA Loper J.; Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
 RL DR [3] EMBL; CP000051; AAY37083.1; -; Genomic DNA.  
 RQ SUBMISSION INFORMATION  
 RQ SEQUENCE  
 RA Feil W.S.; Lindow S.E.;  
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CP000051; AAY37083.1; -; Genomic DNA.  
 RQ SUBMISSION INFORMATION  
 RQ SEQUENCE  
 RA 16.8%; Score 1490.5; DB 2; Length 1682;  
 Best Local Similarity 29.0%; Pred. No. 7-4-72;  
 Matches 477; Conservative 237; Mismatches 715; Indels 217; Gaps 53;  
 QV 14 MSDNNEFPTOANNNTSAVSGGVPDRTGLNIQTLIGHTVNNGNGPTPLTYSPIKANT 73  
 Db 1 MAASTSVHNSNALNFMSCIKSGVPDRTGLYLNISISLPDIQSNLRLGGPRFLDLSQNL 60  
 QV 74 DIERGIGENFGFLSYTDKRSLSLSTGENYKVIEDTKT--VKUQOKUDMLRPEKDLKEN 131  
 61 DSGYGLGMWQMLQSOYDPAOTIISLTGSTGERPTVDTSSNGQMLTMEKKLDEHFKY-LDE 119  
 QV 132 CYRIHKSDIEVLTGFFNNNAFDKVKPGKLNPAGHATYIDMN-FEAQPRMRVYDD 190  
 Db 120 SYRVAKSGLEVIELHLHSGNKRMAPKLIKASGHSLALKHVFNSSTYRASITDQG 179  
 Db 191 GHDIPLLNLYQOSHILTI-----LTLFPGKEGYTRFLRANRQLS---IHNFSLGN 240  
 Db 180 -----QTLIKIARSTTSVELDLHPDGEGTGDTPLARFLMTLAGSDKRVSRITLPT 228  
 QV 241 ENPLTWGCGTPIKGNGTIGQWMTSMTAEGGLKEETVYNNNNQHFFQSANIPVLPVY 300  
 Db 229 ENKASWRPFYGLENGNOQC---VRHVENTPAGSENVY--QDSCHAFSSASGRMPVRVT 283  
 QV 301 LMKVQVPGACQPAIORYSY----TSHNYVGGSNGIN-NNKUDNLGYMTENYGESTS 354  
 Db 284 KHVTDGLNQAKIDVRYTVKGDQORSRNPVAGIPIAMEWDGNDLNUYCLQDQMYCSES 343  
 QV 355 RYKDKEGHDOQIVIERTYNNHILTSCKQOKONGYIQTETAYAATICHNFEDPSOPOL 414  
 Db 344 LWVDNK---AVASIERTFNRPHLQTLTETQNQNLQTTTAYNLOGEHYSRQPNDQCL 399  
 QV 415 PKKKTETRSAD--NSYSEBITEPFDSEGNPLTKVKDKKOKLISSTHMEYPPAGE 472  
 Db 400 PSETTRWORLDAPERTRTETVTDYDSYGNLHVTRADG---IREVSSW--YPAAGG 452  
 QV 473 VDCCPPEPYGTPRFVKLQTPTDSEFKDPKPEPKIQYRSLGS-----QSHV 520  
 Db 453 -DCCPADAELEGFSVRLIKEKVKVKAPSSQKAPLATAVRYQTLALADSELPDMIVPSET 511  
 QV 521 TUKIBERHYSATQOLNSTFOY--NTDKSEL--GRLLKQTECTKGENGKTYSVVHKFTYK 577  
 Db 512 LQOLES---GTCVBLQKULLEYINQPDAPPLHGRTGKQIESLNGNDTT--VTEYSKS 566  
 QV 578 --QDPTLQOSSHSMIT--HDNPFTIHSQVRSYRGRFLFSDTDTKDIVTOMSYDPLGLRUTR 633



Db	1501 GIKLAAAVATATLSVTG--TALQTAELSGHEKRNIGNITILNVSATIIGLAVGAMQLIAK 1558	Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136
Oy	1481 --LATHICPAPFED--GENALKSTSESSRIKGVTSRSDREV 1519	Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594
Db	1559 IPNLAWRAGSYTVTTIDOPPLINSGKTLSBLK----MARQII 1596	Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196
	RESULT 11	Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652
O4594_B COXBU	ID Q55948--COXBU PRELIMINARY; PRT; 709 AA.	Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235
AC Q55948;	DT 01-NOV-1996 (TREMBrel. 01, Created)	Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692
DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)	Orf 709	DR	Orf 709. <i>Coxiella burnetii</i> .
OS	OG Plasmid OPHI.	OC	Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC	Coryliaceae; <i>Coxiella</i> .	OC	"Analysis of the entire nucleotide sequence of the cryptic plasmid
OX	NCBI_TAXID=777;	RN [1]	QPhi from <i>Coxiella burnetti</i> ." ; EUR. J. Epidemiol. 10:413-420 (1995).
RN	NUCLEOTIDE SEQUENCE.	DR	EMBL: X75356; CAAS3129.1; -; Genomic_DNA.
RA	Thiele D., Willems H.; Haas M.; Krauss H.;	DR	PIR: S38241; S38241.
RT	InterPro; IPR006530; YD.	DR	InterPro; IPR006530; YD.
RT	TIGRFAMS; TIGR01643; YD_repeat_2x; 4.	DR	PIR; PFO553; RBS repeat; 8.
KW	Plasmid.	DR	PIR; PFO553; RBS repeat; 8.
SEQUENCE	709 AA; 79416 MW; A377B9A2A72CCF86 CRC64;	Q9	Q98LP2_PSEPK PRELIMINARY; PRT; 1290 AA.
Query Match	13.7%; Score 1214; DB 2; Length 709; Mismatches 287; Conservative 101; Indels 36; Gaps 16; Matches 287; Best Local Similarity 40.8%; Prod. No. 2.2e-57;	Q98LP2	Q98LP2_PSEPK PRELIMINARY; PRT; 1290 AA.
Qy	554 KQTCTKGENGKTYSVVHKPETYKQDPTIQQSHSITHDNPTIHRSOVRSRYGRPSDT 613	AC	Q98LP2
Db	3 RRAEVLTSKEKKYQNQTFAAISSQAEHLIQKIDFTGGDGKITSISREQSRGRLST 62	DT	01-JUN-2003 (TREMBrel. 24, Last sequence update)
Qy	614 DTKDIVTQMSYDQKGLRTML-SGTPYANTLYDVEILANLQDDNRNPPVTTDVG 672	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)
Db	63 DELGNVTOYEVDLGRLTQTHWASNTVASTYRYSLETARGKVAIKTTWDPGEN 122	DE	Hypothetical protein.
Qy	673 QLRNEFDGAGRHSQCLKQSDG----DCKFTYHTQOYDQEQRHHTYSTYDVLTNRQ 727	GN	Orderdlocusname=PP1087;
Db	123 QRRYYDGLGRNUKQERUDKDAVSQKGTWYTHQQYDALGRBESKTRITQDIL----R 177	OS	Pseudomonas putida (strain KT2440).
Qy	728 TDPDKVH----LNSMK--SYDNNQGIAINTHWYSGVSEKTTVDPITLTAKOQNSNNVQ 781	OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Db	178 LDSEVGHGAVSLSTSKVHYDSMGQNHLVFDQYQERSVYDPTTRAT--LQPEGSQK 235	OC	Pseudomonadaceae; Pseudomonas.
Qy	782 TGEKVTTIPSPQQPIQITLDPBAGHQSLHTLTDGMWRVRKETDAGQCYQYDNYR 841	OC	NCBI_TAXID=16048;
Db	236 LQGQLTETVNLAGLPIKUQVYDQDTEOAHSAYEVGDLQQLRKETDELQGQITLYVEHGR 295	RN [1]	YD_repeat_2x; 4.
Qy	842 VIQTLPDGTIVKVKAPSTDTLTDITRNGISLIGOQTFDJSLSRLTOSQDGRWAVY 901	DR	YD_repeat_2x; 4.
Db	296 VTTGTLPEPNTIQKSVAHPSHTSITGTSVNVNNSMGNOTPDSLRLITTSGRTSAFSY 355	DR	YD_repeat_2x; 4.
Qy	902 SAGNDQCPSTVTDQGPHYQVQPELDAAVLOVASNBTITQSYNPTGALIKAV-AEG 960	DR	YD_repeat_2x; 4.
Db	356 ENAS-SVPAVAVTPTGETVSEYELKELGNAVAKKISAPELIQWTQDALTGAMISATOAG 414	DR	YD_repeat_2x; 4.
Qy	961 QSTTPVYPSGRKJNE--MIND---MKKNSYLTWGLGNGYDQGTRQKISRDRHARV 1016	DR	YD_repeat_2x; 4.
Db	415 MIROTYWPSGLKNETSMPPGAQKSTAYTSLAGABQSYTQVYDHEGR 474	DR	YD_repeat_2x; 4.
Qy	1017 QIKSSIKTLYNDLNRHIGSQVTLATGMLTTVERDGLAREIGKLCDSGHTDI 1076	DR	YD_repeat_2x; 4.
Db	475 GIEDNDIKVSLDYDAFGRPTKQGATDKKTVGAVLSTTLTVDDLURREKREISASGQSIVLI 534	DR	YD_repeat_2x; 4.
	RESULT 12	Q98LP2_PSEPK PRELIMINARY; PRT; 1290 AA.	
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		
Oy	1137 YDYNQNIKACHTFADGTEBDAHTEKAMPIDCOLTEFHHTHDPMRNPRLKRDAGRV 1196		
Db	595 .YDALGNMTKQTDFSGG-RNTATVY-SAIDPQLKNUKNDHSYDKPQITLEYDKAGRM 652		
Oy	1197 NTDNHNGNENFTYDTGLRLO--NG--QGSVYGDPLNRLVSQ 1235		
Db	653 R--DRAGT--LXDALSLRQVNGAGAKGQAYDVNTLVSQ 692		
	RESULTS		
Oy	1077 QOSPLKTOOLANRIVKLGVLTQRTBOYSYDSRNLNOYKCDGAKCPTPKYGHISIVTONFT 1136		
Db	535 EQTYORNHLKLERITORGRRTLKRMPAVDSRNLRILIEYTNGEARPODPYGRKAHRPF 594		

Db 234 ANQARWRFITY----EDVIGYLCVRECTPYGGVERFY--QDAGHKPFPSSAARDKQLPR 286  
 Qy 299 VTLMKQVPGAGQPAQIAEYST-SHNYVGGGSNQIWNK-LDLYGLMTEINYGSTERR 356  
 Db 287 VTRHIDRPFQGAAKVKVRVKEYPTFHNGGSSISWSDGDLMLYKVEDTYKSIQI-- 344  
 Qy 357 YKOKEGHDQIVR-TERTVNHYLITSECKQQNGYIQTETAYAALIGENPFDSPSQFOLP 415  
 Db 345 -- QEVRSRVRITRTRPFRLHILTEQATAQDKKLQAFTRYADNAG-NFESQEPYFOLP 400  
 Qy 416 KTKTETTWASADN-SYRSITETTFDESCLNPLTRKIKKTKLISPSTHWETYPAGEV 473  
 Db 401 HDETOQWSLISINPSRQREKKVTRYDSHGNVLTRLPNOVLETW-----WISAEEBEG 453  
 Qy 474 DNCPEPPICGFTFRPK-KIQ- 500  
 Db 454 DE----RGFVRNLKTRTVOPPAMGHGAATLTOAQYRALTPLGS-----YLKOPW 500  
 Qy 454 DE----RGFVRNLKTRTVOPPAMGHGAATLTOAQYRALTPLGS-----YLKOPW 500  
 RQ 512 SLIGSQSHVTLLKEERHYSATOLIN-STLFQ-YNTDKSBLGRLUKQCTCKG-ENGKTY 567  
 Db 501 RLLSET----LSESGSARGTEFLKISKLYQESAERSYGRVQQTSPVSYPGDGSPP 555  
 Qy 568 SVVIRKPTYTKDD--TQOOSHSMITHDNNTIHRQSVRSYGRULFSRDTDKTIVQMSYD 625  
 Db 556 DLTHYGYSLPDDERALQTVKLYGVGDGESKTTIRHALDTGEPFLNLDNGVBYRYYD 615  
 Qy 626 KIGULTRTINSCPYANTLTYPVELNNLQDDNRPPFVITTDNGNLRNEDGAGHV 685  
 Db 616 ALRRVTPREVVAAPGKEPEAEERKVKYFLCAVDENEAQOWAYDVKOVENTHL--LUDLSRPI 672  
 Qy 686 SCLKKDSDG--DGKFYTHTQQYDQEGRHHTSYSDYLTNGROQTDPKVHLSMSKYD 742  
 Db 673 FEBRKKDSATYAGLRLPTYRTRDDELVESTEFDWLGDD--LLELJSHTSYD 725  
 Qy 743 NWGO---IANTHWSYGVSEKXITVDPILT-----ATKOLQSNNSNNVGTGEVITYPSQ 793  
 Db 726 DWGQRYAVLNPDGTVQEE--IDQVASTDGPVHRAWRVE--HSRVSGITETWYNLFE 779  
 Qy 794 OPIOTLFDDEAGHIO-SCHTLTRDGDWDRYKETDAIG--OCTIYQDYNRVIQITLPD 849  
 Db 780 KPVRIERFALDGTSISLQVNDYDGGIGRLRSRQAGGQRVVFYRDADFVREERLAD 839  
 Qy 850 G-TVNRKXAPSPDTLMDRVN---GISLQOPTFGSLRITQSODGGRWVAYTSA 903  
 Db 840 RUVNTYRTYAGHSRNDLPLVSKVGNTESAVALWLBQVFDLRRHTRVATGGRBTFEDP 899  
 Qy 904 GNDQCPSTVITPDGOFIHKQYQPEI-DDAVLQVASNBETQOFSINPVTGALLKAVAEGOS 962  
 Db 900 GERQ-PHWVQAGPDRGTRIEYQRPALGEEPVRLVSLGK-BAKYEDLKNARLTHCEPPGD 957  
 Qy 963 LPTIY-----YPSGRLKMEM--INDMKMSYIWTGLENQYDGTGTTIKISRDPH 1012  
 Db 958 ENSGYTLDRSHFLSNGEVKRERSTVDGEAFFSMTIDSFRSRRLAAYDVQQTQIXDFDDV 1017  
 Qy 1013 GRVTOQIKDS-----IKTTNYYDINRHGSQVNDLATCHMLTT 1052  
 Db 1018 GRLEKTYLHAEPEKPTVYRLQRAPAROLLINESTPGYDROGRMASITTTDASTGHALATL 1077  
 Qy 1053 VEFDSLAREGRKLCDSGSGHLDIQQSWLKTQQLANRIVK-----LNGVLOREOVS 1104  
 Db 1078 LEYDFDRETRTTE-DFGDPTWQTLAQDYBDFDCLSKRILGRPKGSDESOATLRRHETO 1136  
 Qy 1105 YDSRNRLNQYKQDGAFCPTKRYGHSTVTONTFTYDIGNITACHTFADGETHATFKFAN 1164  
 Db 1137 YDRRGRLQIYTCDGFBAPVDPSSQTIARQIFGGFDGLDNILSITVTRPDGSWQRTYEFKN 1196  
 Qy 1165 PTDPCQ-----LTVFHHTHPMPD-----NIRLKQDKAGRUYINTDHNHN 1204  
 Db 1197 -SDPAQMSRITPPDALLEYTDVHDHOLEBLFRKIVKVDQPLAIDLHYDNGCNL--ISDEQGR 1253  
 Qy 1205 TENFYDTLGRL-----QNGGGSVYQDPLNL 1232  
 Db 1254 V--LTVGDNLNLLRVETPDGSRRCRNYDOPENIL 1284

RESULT 13  
 Q883W6 PSESM  
 ID 0883W6 PSESM PRELIMINARY; PRT; 982 AA.  
 AC 0883W6\_PSESM  
 DT 01-JUN-2003 (TREMBrel. 24, Created)  
 DT 01-JUN-2003 (TREMBrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)  
 DE YD repeat protein.  
 GN OrderdlocusName=PSPT02231;  
 OS Pseudomonas syringae (pv. tomato).  
 OC Pseudomonadaceae; Proteobacteria; Gammaproteobacteria; Pseudomonales;  
 OC Pseudomonadaceae; Pseudomonas.  
 RN NCBI\_TaxID:323;  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=DC3000;  
 RK MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;  
 RA Buell C.R., Joardar V., Lindberg M., Selengut J., Paulsen I.T.,  
 RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,  
 RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,  
 RA Nelson W.C., Davidsen T.M., Zafar N., Zhou L., Liu J., Yuan Q.,  
 RA Khouri H.M., Pedersen N.B., Tran B., Russell D., Barry K.J.,  
 RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Accenzo M.,  
 RA Dang W.-L., Ramos A.R., Alfonso J.R., Carrinholo S., Chatterjee A.K.,  
 RA Delaney T.P., Lazearowitz G., Martin G.B., Schneider D.J., Tang X.,  
 RA Bender C.L., White O., Fraser C.M., Collier A.;  
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
 Pseudomonas syringae pv. tomato DC3000."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).  
 DR EMBL; AB016853; AAC055747.1; -; Genomic\_DNA.  
 DR PTCTO231; PSPT0231; -.  
 DR InterPro; IPR006530; YD.  
 DR Pfam; PF00553; RHS repeat; 6.  
 DR TIGRFAMs; TIGR01643; YD repeat\_2x; 2.  
 KW Complete proteome:  
 SQ SEQUENCE 982 AA; 106800 MW; 00E7CA6B93819B78 CRC64;  
 Query Match 10.5%; Score 932.5; DB 2; Length 982;  
 Best Local Similarity 29.7%; Prod. No. 7.9e-42%;  
 Matches 385; Conservative 142; Mismatches 357; Gaps 33;  
 Qy 642 ANTLTYDVELNILQDDNRPFPVITTDNGNQLRNEPFGAGRAVHSQCLKDSDGKFTI 701  
 Db 31 SSTLGYRD-----DWGR--CCUTTDN-VQTYEYSPPIGSVPHK-----GPIQK 74  
 Qy 702 HTQOYDEOCRHHTSYSTSYLTLNRRQQTDDPKVHLSMSYSYDNGQIANTHWSYGVSEKIT 761  
 Db 75 WKQSGDPPEQ-----ISGRSET-----WLNL-----FOKPDPR 103  
 Qy 762 VDPITLTATKQLOSNNSNNVGTGEVUTYPSOOPIQITLFDAGHLSOCHTLTRDGMVRV 821  
 Db 104 -----TLLACKTGASRTHNSRSRNLTT-----TSQEL-----SRQTFLYDGLGARC 144  
 Qy 822 RKETDAISGCTIVQDNNRVIQITLPGTIVKARYAFTSDPLTIDRV--NGIS-- 875  
 Db 145 TEORDALOOSTLFSYDNNSRMVSSTLADGSVINRSYAPQSSSLATMLRWHONGTTKV 204  
 Qy 876 LGQTFDGIISRLTQSQDGGRVWAWTYSAGNDQCPSTVTPDGQFPHQYQPELDAVQV 935  
 Db 205 AGTQKEDGSLERVWQTKDGRVQFNYDQEMO-PRSRITAGLONINPTVTRALDQISS 263  
 Qy 936 ASNITQOFSYNPVYTGALVAGQSL-PIPIVPSGRJAKMENINDMKNSYLV-----T 989  
 Db 264 TAPPETAKEDYKTSARLIEATWFGTTRYDHNQLTGETWDNL-LGQAWETRHOSS 321  
 Qy 990 LRGLENGYD-----TGTQIKISRDTGTRVQIKDSIKTLYDMMRHSQVTDIA 1044  
 Db 322 LLGPRKTRDILKKGEEAAABTRYDVTGTRIPRINOSHNLRTTIDVQGQCKVATELQ 381  
 Qy 1045 TGHMLTTWEFDGLAREGRKLCDSGSGHLDIQQSWLKTQQLANRIVLQRTQVS 1104

Db	Best Local Similarity	Local Similarity	Matches	Conservative	Pred.	No.	8-1e-41
OY	1105	YDSRNRLNQYKCGAECPTDKYGHISIVTONFTDYGNTTACHTTFADGTEHATKFAN	1164				
Db	442	YDPRGLITUVNLQSLPFLDELQRBMTRQIFSDELONLTCQTRFTGTTSERAKIGS	501				
OY	1165	P-----TPCQLTEVHHTPDMPEINTRIYKDGGRVINTDNEGNENTPYTDIGR-----	1216				
Db	502	PGDDKKHCKRCOLISIAYTTPRKTPDPTFSYDANG--QLKDEHGN--SLHYDQSRLQV	557				
OY	1217	-QNGQG--SVYGDPLNRLVQSQTDTLDCB--YRETMVNEVRNGEMIRLRTGEVII	1271				
Db	558	AETGGAPPSQYRQDNGQVAT--DGNESRDPYEGHOLSTVOEDORTOYHLGROPL	616				
OY	1329	PVGWTGNGVSYDPTLMRFPMDSPSPFGAGGINPVSYCJDPTNSDPSH-----	1383				
Db	677	AANGWYLLGNGYRANYNPPLMRSPDPLSPFABGGVNPTCYLGNPLALRDPTGHDASQ	736				
OY	1384	-----LSWQ-----AVTGIGMIA-----GULLTAGGMIAAAGGI-----	1416				
Db	737	TGRLRPPEGALPQMGCDIMGWGVGWTGVAAITLTGTTPTVGVTVIGI	796				
OY	1417	-----AMAIASSTTALAFG---ALSVTSDITSVAGLEDASKPKASSILGWS---	1462				
Db	797	SMTASAAAVSTVSTGALLVGTALTAASTANTVAVNN----DOTAGEVGWLGIAV	851				
OY	1463	---MGMGIAAGLAEASKIGKTLA-----THIGARADEGENALKTSSESR	1505				
Db	852	PVGLVGFAGAVVARAVAAKVAANAGTIVRSVSRIKG-LAAGGARTISSAASSAR	909				
RESULT 14							
ID	087V6_G_PSES	PSES PRELIMINARY;	PRT:	1669 AA.			
AC	087V6_G						
DT	01-JUN-2003	(TREMBLER, 24, Last sequence update)					
DT	01-OCT-2003	(TREMBLER, 25, Last annotation update)					
DE	VD repeat protein.						
GN	Ordereridcuname=PSPT04970;						
OS	Pseudomonas syringae (pv. tomato).						
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.						
OX	NCBI_TaxID:323;						
RN	[1]	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].					
RC	STRAIN=DC3000;						
RX	MBLINS-22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;						
RA	Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T., Gwin M.L., Dodson R.J., Daboy R.T., Durkin A.S., Kolonay J.F., Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Hart D.H., Nelson W.C., Davidsen T.M., Zafar M., Zhou L., Liu J., Yuan Q., Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J., Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Aszenzo M., Deng W.-L., Ramas A.R., Cartinhour S., Chatterjee A.K., Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L., White O., Fraser C.M., Collier A.; "The complete genome sequence of the Colletotrichopsis and tomato pathogen pseudomonas syringae pv. tomato DC3000," Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).						
RA	EBML: AE016053; AAO58398.1; -; Genomic_DNA.						
DR	INTERPRO: IPR06530; YD_RHS repeat; 5.						
DR	TIGRFAMS; TIGR01643; YD_repeat_2x; 2.						
KW	complete proteome.						
SQ	SEQUENCE 1669 AA, 184268 MW, BC668B6C1E3FB7EE CRC64;						
OY	1004	IOKLSRDTPIGRVQIKOSSIKITLNYDDLNHRIGS---QVTDLATEGHMLTIVWEFDGLN	1059				
Db	825	FPGCS---ASTVTPRDKTLESAVSADCHATATINSKVKHTPVISOPNLVATVDPV	881				
OY	951	-----GALKAVAGQGSIPTPIYY--PSGRKMKENINDMKKMSVWLTGLENTYDGT	1003				
Db	933	QMRVFRDYLDRVVRV---ILSELHYHVLWSAFGPLORTVWNQASGERLDRVPTWDAG	988				
OY	1060	REGRKLKDSSGHTLDIQQSMWKTQQLANRIVLNGVQORTGCVSYSPSRNRYQKC--D	1117				

Db 989 QETAREYTLANKKPLALNTSYLNGQVSSTLTREGVILQRTGFSYDARDRLNSYECTD 1048  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1118 GACPTDKYGHISVTQNTFTIDYENITIACHTADGDFHATKE--ANPPDCQUL-TE 1173  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1049 VAWPQDODAGKSLSKQSSTGYSYGYDELAHLNLSCSSTADGSTCIOTYTYDWTWNPTRLSVKE 1108  
 Db 1174 VHH-THPDMPDNIRKYDKAGRVRVINITDNGNTENFTDTLGL--ONGQESV--YGY 1226  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1109 LRGSSQTSTQTATLAYDANGN--QTIBSGRT--LAVPLSQLASVNDGAKLTRY 1164  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1227 DPLURLVSOKTDI--DCELYRBTMLNEV--RNGENMURLT-----GETI 1271  
 Db 1165 DAFGRLLSOGIVATKHTBELLYDGTOLIGERAWFDDANEKFPLSFEDWVQTCIGEV- 1223  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1272 AQQRASKVLTGDSQSQSVLTSKQ---NLSQEA-SAYGHKSTANDASI-LGTYNE 1325  
 Db 1224 |----RSDFVL--.TDPGSGVVGFSAHGDTAGVKHPLGTYPG--ESTNLSDGCRIGENSE 1275  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1326 RADRVGVTHLNGRSYDPTLMRFPHTPSLSPRGAGGINPNPYCYCLGPINRSDPSGL- 1384  
 Db 1276 RIDPVLGWTHLGNQYRTSPAQRHILQPSWPSPGAGGINNTAVCAGPVNLIDPSGM 1335  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1385 --SHQA-----WTG I---GMIAGIAGLTTIANGMAIAAGGI 1416  
 Db 1336 ISRGAGASNMTSLTKAQLQSTSQPQLGHFWRGAVLNASVAVAGVLMVPLTGGSSLGFAAGV 1395  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1417 AAATASTSTALAFGALSUTSDITSIVSAGLEDASP-----ASSLGWVWGMGAG 1469  
 Db 1396 LA-----TTLAVASAGL---DIASYV---LEDVNPELKRGTAATALGFLSNAPFRAG 1443  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1470 L-----AESAKGGTAKATHLGF--AEGDENALLKTSSESSKIK 1507  
 Db 1444 LRIGGRLRWTSRRTGRLSRVSEIFTKPKPLKLITGKLVKEG--MSYGASNYTR 1499  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1508 WGVTRSLDREBIVRNIEBEGQVTKDHSGRYTNFMGKGEQATLVHGKDGD---FLYHTEGNK 1563  
 Db 1500 W-----APGKUDKF-EKLENALFMTWDQKGKRVTFMAH-GVK 1534  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 1564 -----HNGKGPYI-RHTPBLQVYDYLKDNNIVDILTQGGDKPVHILSCYGS 1607  
 Db 1535 PDDVGDAMMAEEVEYSGSGPSPYSGKSFSAVSPFDLKSKK-VDLNK-YEVKRIMCHSAD 1591  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1592 GGEKSFAATESKLTINKPKVSKY 1612

RESULT 15

Q45905 COXBU  
 ID Q45905\_COXBU PRELIMINARY; PRT; 505 AA.  
 AC Q45905;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Orf 505 protein.  
 GN Name=orf 505;  
 OS CoxIELLA burnetii.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;  
 OC CoxIELLaceae; CoxIELla.  
 OC COXBU  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE  
 RX MEDLINE=97294470; PubMed=9150226;  
 RA Willens H., Ritter M., Jager C., Thiele D.;  
 RT Plasmid-homologous Sequences in the chromosome of plasmidless  
 CoxIELLA burnetii Saurry Q217;"  
 RL J. Bacteriol. 179:3293-2297(1997).  
 DR EMBL; X932204; CAA63682.1; -; Genomic DNA.  
 SQ SEQUENCE 505 AA; BE60E4C307CABD0B CRC64;

Query Match 10.4%; Score 919.5; DB 2; Length 505;  
 Best local similarity 41.1%; Pct. No. 1.4e-41;  
 Matches 217; Conservative 75; Mismatches 195; Indels 41; Gaps 16;

Search completed: January 30, 2006, 09:49:02  
 Job time : 141.373 secs

18 NEP-FTQANNFTSAVSGCVDPRTGLANLQITLGHIVNGNIGPLTPILSISPLNKDII- 75  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 2 NELPYTQATNFISAVOCVDPRTGLIVNVMLAELTGNDNLGPDLFTLNTSPSLSNIC 61  
 Db 76 GSGIGENGFLGSLYDRKNSLSSISTGENGKVLTEDKTWQOKGLDNRPEKOLKENCRI 135  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 62 GFGIGCSVGISYIDKNNKLILSSSERYKTEDWNSDVYVRQCKINNFKFEK--IKNGYI 119  
 Db 136 IHSGSDTEVLITGFNNNAEFLKPKKLUNPAHAIYIDWNNEATOPRNRIVYDLDGHID 195  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 120 KYRKQKTYLYKGDNF--LPQKTFSPGLPKLUSWENRQSYVNITKIEDAD--V 172  
 Db 196 LILNLEYQCGLIKITLTLEPGQKEGYRTERFLRNLQMSIHNFSIGNENPLTMWSFGYTPIGK 255  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 173 LICKDQYPSDWMARITFEGKIESYTFQDFVNEYLYWTNGESTSRE--LWMSFNYDVG 230  
 Db 256 NGTLGQWNTSMWARGKAKTETVYNNNGQHHRPQASAMPVLPVYVTLMKQVPGAGQPA1QA 315  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 231 GNFT--.TQVQKSPGTLGETVNTQAGM--RFPDESGKPALPSVINYRQSPGMQPDIVK 285  
 Db 316 EYTSYHNVG-GGSNGIWNNGLNLXG-LWTEYNGSTESRKYDKEGHGDLQVIRITY 373  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 286 EYXTASYLTYLGKGAISLGRANEDBNITVWMDYTSSTEKLIVDNR--LVSISRY 341  
 Db 374 NYTHLILTECKQOONGYIQTTEAYAATIGHNFDQSOPQSPQDPLKTKETWR-SADNSYSE 432  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 342 NSTYLLSSETTRONSCREVIVETDYAKRGLSDKQPKQPKQFQPLKEKWTRENKQDCESE 401  
 Db 433 ITTEPFDPSGNPLTKVDKKTKQKITSBSTMHYYPPGEVDN---CPPEPYGFTREVK 488  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 402 ITTTFDPEGNLTIKIED-----GTXTEVYIYDSKGETDKGIVLCPPEPGFVREVK 454  
 Db 499 KIQTQPUSEPKDPEKETIQYRS---LIGSOSHVWIKIERHYSAT 532  
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 455 TQIWTIPANSEFY-APOQDVTAAQYPCTAGSISLYAVLQNKKHFAVT 501